# *Review Essays*

Psychology of Music

*Psychology of Music*

Society for Education, Music
and Psychology Research
vol 32(1): 105–129 [0305-7356
(200401) 32:1; 105–129]
10.1176305735604039286
www.sagepublications.com

D. TEMPERLEY, *The Cognition of Basic Musical Structures*. Cambridge, MA: MIT Press, 2001. 404 pp. ISBN: 0–262–20134–8 (hbk) $45 (£30.95)

*The Cognition of Basic Musical Structures* is an ambitious work by David Temperley. The author is a music theorist, composer and performer with expertise also in music computation. He aims to describe musical infrastructure, the 'ubiquitous . . . means to an end' that makes music work, in the same way that 'roads, power lines, water mains and so on . . . needed for life and business . . . facilitate . . . homes, schools, showers, VCRs' (p. 3). In other words, Temperley is attending to a necessary but ordinary aspect of music which music theorists, more interested in new interpretative insights, would typically assume or ignore.

The book has two parts. The first, entitled 'Six Preference Rule Systems', directs attention to six music structures identified by the author as basic. They are meter, melodic phrase, counterpoint, tonal pitch-class representation or pitch spelling, harmony and key. The second part of the book entitled 'Extensions and Implications' addresses issues posed by ambiguity, rock music, meter and grouping in African music, style, composition and performance. In this review essay, I briefly summarize Temperley's views and then provide a general critique, all from a music-cognitive perspective.

## *Part I: Six Preference Rule Systems*

Temperley explains his approach in chapter 1. In the first part of the book, he outlines rules that describe how he, as a listener, hears each of the basic music structures in examples of music from the 18th and 19th century (the common practice period). He refers to these rules as preference rules, a term popularized by Lerdahl and Jackendoff (1983) in the *Generative Theory of Tonal Music* (*GTTM*). (Lerdahl was in fact Temperley's doctoral thesis advisor.) Temperley identifies a basic music structure, describes it, then develops the computer algorithm to analyze it (in collaboration with computer scientist Daniel Sleator), and then tests the algorithm.[1] Algorithms that analyze meter, pitch-spelling, harmony and key are tested against analyses of music

excerpts in the Kostka and Payne (1995) theory workbook. This corpus provides 46 excerpts of eight or more measures for which harmonic and key symbols were published. The second part of the book extends the discussion to other broad issues.

Chapter 2 begins with metrical structure. From the start, a reliance on introspection is apparent, as opposed to an empirical approach.

> There is a general consensus that meter is an important part of music cognition. However, in modeling the perception of meter, how do we determine what metrical structure is actually heard? In this chapter, as elsewhere in the book, I will rely on my own intuitions about the metrical structures for pieces, and assume that other listeners would generally agree . . . In most cases the metrical structure indicated by the notation for a piece agrees with my own intuitions, and I think with those of most other listeners as well. (p. 26)

This introspectionist approach characterizes the book. There is probably no better extensive systematic music introspection available than Temperley's but when unaccompanied by empirical support it may upset contemporary experimental psychologists. They generally acknowledge insights from introspection but prefer additional experimental validation involving human listeners. Temperley does not provide this validation. However, readers who are disturbed by the introspection at the beginning of a chapter may find comfort at the chapter's end when the computer algorithms often nicely model Temperley's intuitions. It could be argued that such success with the algorithms earns Temperley the right to avoid collecting data from human listeners.

Temperley derives a preference rule system for meter and then, through computation, implements its ability to choose a time signature. Tests of the effectiveness of the algorithms are then conducted on both quantized and unquantized MIDI data derived respectively from direct entry of the notation of the Kostka-Payne examples and from an actual performance of the Kostka-Payne examples. In both cases, the raw data produce a 'piano-roll' which as Temperley admits overlooks the complex psychological question of how the listener determines a pitch category in the first place (cf. Meredith, 2002). The results of these analyses are successful. For the quantized data, the five levels of metrical analysis (two below and above the tactus, i.e. what the listener regards as the beat) are matched correctly at least 86 percent of the time. The success rate is still above 70 percent for the unquantized data. Temperley notes that parallelism plays an important part in meter identification but it is too hard a problem for him to worry about here (p. 51); he has since published a paper in *Music Perception* on this aspect (Temperley and Bartlette, 2002). The same avoidance of hard problems arises in the next section on segmentation, and here it really limits the relevance of the analysis.

In chapter 3, segmentation, or grouping according to phrase structure, seems more of a challenge to Temperley than meter was. To exemplify the problem, he considers a particular location in a passage from a Bach

Invention (pp. 64–5). He states it is impossible to decide whether a segment should begin at this point. He then claims 'it seems reckless to attempt a computational model of an aspect of perception in which it is so often unclear what the "correct" analysis would be' (p. 64). So, instead of that hard problem, he decides to model segmentation within a simpler corpus. 'Rather than attempt this, I will address a smaller problem [of melodic phrase structure]' (p. 64). He chooses to model examples from the Essen corpus of monophonic folk melodies representing 20 European ethnic groups (Schaffrath, 1995), some of which may have little similarity to music of the common practice period. The selection chosen contains 250 boundaries indicated by phrasing in the published corpus.

Having derived three preference rules, Temperley's first pass with an algorithm incorrectly predicts 115 boundaries (i.e. *not* indicated in the corpus of 250 boundaries) and misses 56 boundaries (i.e. that are indicated). By changing the phrase length parameter of his algorithm, however, these numbers reduce the incorrect additional boundaries from 115 to 66 and increase the missed boundaries from 56 to 63, respectively (p. 74). Thus the program on this last test correctly identified 75.5 percent of the boundaries and conversely missed 24.5 percent. According to Temperley, in assessing the model, consideration of only the 63 false negatives (misses) and not the 66 false positives 'seems reasonable, given that the number of false positives and negatives were roughly the same' (p. 74). In other words, he argues that because the false positives balance the negatives, one can ignore the false positives. In fact, explaining boundaries that are missed by a theory but 'heard' by the listener may be quite a different problem from explaining boundaries predicted by the theory but 'unheard' by the listener.[2]

Where would listeners place boundaries? Without considering data from listeners on boundary formation, Temperley chooses as correct the notated phrasing of the Essen corpus of folk songs. The fact that the Essen corpus is not always representative of common practice music raises the question of whether it reflects the real experience of listening to and performing phrase segmentation of common practice music. Eventually, other researchers may test the algorithms against human listening data and provide closure to these issues.

In chapter 4, Temperley defines a contrapuntal analysis as 'a set of streams, subject to various constraints' (p. 97), and represents a stream on a graphic 'piano-roll' matrix derived from the score. After outlining four well-formedness rules, four preference rules are defined: pitch proximity ('a stream must consist of a set of temporally contiguous squares on the plane', p. 97), new stream, white square rule, and collision rule. Bach's notated fugues provide a good test of the model because the streams are clearly delineated. Still, many issues of ambiguity cannot be addressed satisfactorily, as Temperley admits.

In the next chapter (ch. 5), Temperley introduces a controversial aspect of

the book concerning pitch spelling and tonal-pitch-class representation. He feels strongly about appropriately distinguishing enharmonic spellings. Believing that these different spellings have psychological validity, he distinguishes between tonal pitch class (TPC) and neutral pitch class (NPC). The former refers to the music-theoretic distinction such as between A♯ and B♭, B♯ and C, and C𝄪 and D. For example, when spaced out in the 'line of fifths' B♯ and C are separated by 12 units. On the cycle of fifths, however, they overlap. Temperley points out that NPCs are used in cognitive psychology (e.g. probe-tone profile of Krumhansl) but TPCs are used in music notation and traditional tonal theory. Temperley believes that the distinctions are relevant for highly trained musicians but perhaps not for other listeners (p. 123). He remarks that surprisingly there has been no psychological research on the problem: 'the cognitive reality of spelling distinctions simply has not been explored' (p. 123). Yet he chooses to base his chapter and preference rules on TPC as a prerequisite to harmonic analysis. Temperley offers an example of intervals of the same size having different spellings (minor third vs augmented second, p. 119) which do sound different, but their contexts are also completely different. Whether the differences are attributable to TPC demands empirical test. Instead, he resorts to emotional rhetoric: 'The placid, stable minor third in figure 5.6a has quite a different effect from the restless, yearning augmented second in figure 5.6b' (p. 119). The rhetoric may not convince everyone. Experiments are needed (cf. techniques for determining connotations of musical intervals by Costa et al., 2000; and for the integration of all interval information in a melodic configuration, Cuddy and Cohen, 1976). Chapter 5 could inspire a wave of research on the perception of enharmonicity in answer to the question: Are accidentals in music notation based on orthographic simplicity or on musical grounds?

Temperley (1999) had already published his views (much of them verbatim) on pitch class and the line of fifths. His examples convince him but are there other reasons for the use of particular spellings? Compare the sheer simplicity of writing D as opposed to C𝄪 (see also Marvin and Brinkman, 2000). Even people unschooled in music would agree that accidentals add complexity. To avoid this unnecessary complexity, Temperley allows for transposition, claiming that people hear relatively and to argue otherwise is to argue for general capacities for absolute pitch. 'For one thing, treating the line-of-fifths space as absolute would assume listeners with absolute pitch' (p. 127). There may be more psychological evidence of the general capacity for absolute pitch (Halpern, 1989; Levitin, 1994; Terhardt and Seewann, 1983) than there is for the need to preserve distinctions between A♭ and G♯. Temperley overlooks much scholarly debate when he states: 'There is no musically important reason for notating a piece in C major rather than B♯ major; it is simply a matter of notational convenience' (p. 127). This ignores discussions of why a composer chooses a particular key for a piece. Yet, in fairness to Temperley, through what might seem as convoluted reasoning and

contrived preference rules, he creates a most impressive algorithm that identifies the 'correct' tonal pitch class 98.8 percent of the time. Credit must be given for the definite capture of something. The achievement, however, requires positing a Center of Gravity (COG) of the tonal pitch class set for a given key. This wonderful acronym, however, seems wasted on a construct that lacks cognitive reality. For the key of C, this COG is D. The note D, as a non-triad diatonic tone in the key of C is not typically considered psychologically significant. The pitch chroma of C is the psychological COG – this centrality of C (not D) is the essential phenomenon of tonality. So, in spite of positing the psychological significance of his algorithms, Temperley here posits a construct that may have little relation to anything psychological.

One valuable aspect of Temperley's pitch class proposal is the insistence on the importance of the cycle of fifths representation. Translating this cycle into his construct of the boundless line of fifths however (e.g. Figures 5.5, p. 118, and 5.11, p. 127) may add complexity which is irrelevant from a psychological standpoint.[3]

Temperley begins chapter 6 on harmony with '[t]he importance of harmony, is indicated best, perhaps, by the amount of time devoted to it in undergraduate music education' (p. 135). Perhaps not. Does it follow then that importance of the psychology of music is indicated by the little time devoted to it on the typical undergraduate music curriculum? A curriculum depends on many factors including a method for teaching and examining a body of knowledge and the availability of a qualified teacher. Of more theoretical significance, Temperley daringly states that what is important in harmonic analysis is root not chord identification (p. 138). This emphasis minimizes the role of key identification, generally the first principle for harmony analysis. Traditionally, the key must be identified before chords can be analyzed with respect to their function in the key.

Thus, challenging convention, chapter 7 on key identification *follows* chapter 6 on harmonic analysis. Temperley rejects early models of key analysis that rely on the last notes of the piece because 'we are surely capable of identifying the key well before the melody ends' (pp. 169–70).[4] To identify key (TPC), he adopts Krumhansl and Schmuckler's (in Krumhansl, 1990) algorithm that uses empirically based values as a template. Following a process of trial and error, to improve performance of the model, he arbitrarily increases emphasis on the fourth and seventh diatonic notes beyond their values that are based on Krumhansl's empirical data. Temperley's acceptance of the Krumhansl approach to key indicates the compatibility of his approach with obtaining empirical data. The larger scope of Temperley's as compared to Krumhansl's (1990) book makes acquisition of such data more daunting, and then, perhaps justifiably so, Temperley makes no attempt to acquire empirical data other than in borrowing the key-finding algorithm.

Chapter 7 completes Part I. A word is in order about the implementation of the algorithms associated with each basic music structure discussed in

chapters 1 to 7. The algorithms use a dynamic programming and a best-so-far approach. The process is sequential. Each musical event in sequence provides potential choices. To summarize, in an analysis of boundaries by the model, at each note, a boundary could be imposed. In an analysis of meter, each beat was identified at one of five different accent levels and the time signature was assumed on the basis of the middle level. In analysis of roots, one of 12 possible root notes was assigned. In analysis of key, one of the major or minor keys was assigned. For each of these, there were both numerical penalties and benefits associated with making a change to the status quo of the analysis as the piece unfolded. Computation of all possible outcomes at any particular point in the piece enabled a judgement as to the best analytic choice so far. Having made that choice it was not necessary to revisit that choice point again. When the final choice point was assigned at the last note of the piece, the entire sequence of choices was in fact the analysis. Examination of the analysis however might reveal several inefficient zig-zags (e.g. switches of root or key). In point of fact, this best-so-far heuristic is only a means to an end, and not necessarily the best end. Hence, for purposes of finding the most sensible analysis, intuitive revisions of the best-so-far outputs were allowed. The resultant analysis is akin to a theory of listening for that basic structure as represented in that example of music.

## *Part II: Extensions and Implications*

In the first of the chapters of Part II (ch. 8), Temperley discusses ambiguities of meter, harmony, pitch spelling, grouping and counterpoint. He raises issues about whether the brain can entertain more than one interpretation at a time (p. 222). Chapter 10 on African rhythms is the only one in which he does not rely on his own introspection but rather those of experts in this field. The analysis leads to a notion of relative weighting of rules such that the total weight is still the same for western and the non-western style (p. 289). Reference to the information processing approach and fixed mental capacity, originating with George Miller (1956), would have provided a cognitive argument for this postulate. Chapter 11 on style, composition and performance develops a scheme for determining the degree of surprise value (or, conversely, degree of boredom) of a particular piece within a style, in accordance with its adherence to the preference rule characteristics for that style. This notion of rule adherence resembles Berlyne's (1971: 193) theories of hedonic value, that acknowledged a debt to the Wundt curve, although Temperley provides no reference to this.

The consideration of rock music (ch. 9) adds another dimension to a text that primarily focuses on music of the common-practice period. It would have been appreciated if Temperley had defined what he meant by rock. Why did he select particular pieces as examples and not others? In identifying how rock differs from common-practice music, he emphasizes that rock typically

communicated through an oral and performance tradition as compared to the notational origins of common-practice music. Yet, scores are published of rock music (e.g. by the Beatles), but these seem not to have been consulted. Instead, Temperley provides his personal transcriptions. Temperley suggests that appreciation of rock music demands that the listener create both a surface and a deep structure of the unfolding piece. The deep structure arises by a shifting forward operation carried out to cancel syncopation. This places long tones on strong beats and in theory allows for analysis of meter and harmony using algorithms that worked for common-practice music. As for tonality, Temperley, acknowledging Moore (1993), argues that rock music makes more use of the Mixolydian and Dorian modes than common-practice music does, and that the same rock piece may use all four common modes. Thus a wide range of the line of fifths represents the set of pitches used by a rock piece (see the 'supermode', Figure 9.22, p. 260). Another way of putting this is that for a C tonic, F♯ and C♯ are uncommon. It would appear then, that for such a key, modulation to the dominant could not be prepared through a V-I cadence in the new key (which would require the F♯). Why this is so is not addressed. Nor is the question of chord progression, which Temperley argues differs from that of common-practice music, a structural aspect almost entirely overlooked in the book.

The final chapter (ch. 12) titled 'Functions of the Infrastructure' is an attempt at synthesis. Here several prominent aspects of music, previously neglected, are also addressed, but somewhat superficially and the chapter becomes a bit of a catch-all. Meaning is defined in terms of its function not in terms of associations brought to mind by the music although the concept of association is later introduced. Why then not include it in the definition of meaning to begin with? Motives are described in terms of repetition. Temperley concludes that motivic structure facilitates memory but seems to overlook the question of what kinds of similarities the brain can pick up. He discusses and approves Deutsch and Feroe's (1981) model for representation of tonal sequences, but worries about its failure to consider rhythmic information. Here he postulates that repetitive rhythmic patterns would be easier to remember than those lacking repetition. Some work along this line has been conducted by Royer and Garner (1970), and Temperley does acknowledge other work by Povel and Essens (1985), so exactly what Temperley is looking for when he says the work 'has not been done, to the best of my knowledge' (p. 330) is unclear. It seems that Garner's (1974) *Processing of Information and Structure* previously made the points that Temperley is trying to make. A few pages later, however, Temperley states: 'It is undoubtedly true that rhythmic repetition aids encoding independently of grouping' (p. 334). It is unclear to what data he is referring. The terminology is confusing since rhythm involves grouping. In this section he also attempts to show how the different kinds of structure that he has previously isolated can be a basis for motivic similarity. He raises the question of the effect of

their interaction. He concludes this section with the hypothesis that 'It may be that an analysis that permits a "rich" motivic structure (one that contains many coindexed segments, and can thus be very efficiently encoded) is preferred over one that does not' (p. 336).

Temperley introduces a theory of 'metrical parallelism' nicely illustrated with a comparison of two repeated phrases, one with equal metrical relations, the other without. Intuitively, the melodic similarity is evident only in the former case. This leads to the notion that music processes 'can be seen as a search for order and pattern' (p. 334). He points out the similarity between musical listening with other mental activities involving guessing (jigsaw puzzles, games like charades, and 20 questions). The appeal he argues is in finding order amid complexity. This then is an aesthetic theory having antecedents in writers such as Koestler (1989[1964]), and Platt (1961). A review of these and related concepts can be found in Berlyne (1971, ch. 10). Temperley fails, however, to acknowledge any of these prior ideas. With specific reference to expectation of musical patterns (p. 334), it is not clear why Narmour's (1990) Implication–Realization model is not referred to here because it is founded on this concept. Narmour's mentor Leonard Meyer, however, is acknowledged in a footnote.

In this same last chapter, a concept of musical schema based on a definition from psychology 'a cluster of features' (p. 336) is established. Temperley gives no source for this definition, nor any reference to the vast psychological literature on this concept, a concept which often emphasizes active mental processes and past experience, aspects uncaptured by Temperley's treatment.[5] He refers to 'musical patterns that occur in many pieces, and are thus the property of the common-practice style as a whole. Such a pattern is frequently called a "schema"' (p. 336). It is however the mental representation that is the schema not the pattern. He then refers to the V-I (perfect) cadence as the most important schema. Later he states that psychological evidence of recognition of perfect cadences by listeners leads to the assumption that listeners 'must be performing harmonic analysis as well' (p. 339). The general nature of this assumption is bewildering because Temperley has earlier proposed that harmonic analysis is accomplished via TPC categorization, root analysis, key determination, and then finally a chord function analysis. It seemed earlier we were already far beyond this assumption that listeners perform harmonic analysis when listening to music.

Temperley continues to add more new concepts: musical energy and musical speed. In a subsection of Chapter 12 titled 'Arbitrariness', Temperley attempts to deal with the origins of musical meanings. Occasionally, he notes, associations are non-arbitrary. His example: 'March music symbolizes the military because it is actually somewhat similar to military music' (p. 348) seems like circular logic. After a consideration of four hypotheses, his conclusion seems to ignore much work on the topic of music and language, even that of which he is well aware: 'The fact that there is nothing in music

comparable to the arbitrary relation found between form and meaning in words is an important difference between music and language which has not been fully appreciated' (p. 348). In the penultimate subsection on 'Musical Details and Recomposition', Temperley tries to show, by means of what he calls an experiment, that the preference rule approach can predict the 'consequences of microscopic compositional decisions and changes' (p. 354). In the final subsection, titled the 'Power of Common-Practice Music', Temperley describes a number of characteristics of music such as 'complex emotional associations' and motives that 'rely crucially on several aspects of the infrastructure' (p. 354). He states:

> As other authors have noted, motives often serve as 'agents,' entities with feelings, desires, and capacity for action. Combined with the expressive powers of the tonal system, this allows composers to create complex dramatic narratives. (p. 354)

At least the reader is assured that 'other authors have noted' them, but anthropomorphism (i.e. inanimate objects, in this case motives that have 'feelings' and 'desires') in these final paragraphs comes as a shock. Our original bargain with Temperley was to accept his introspection provided that he could model it. Now he divulges that musical tones for him (and for other unnamed authorities) actually have a will of their own, a will that would challenge the modeling approach.

The reader might have been convinced until now that Temperley could model important basic aspects of its structure. Temperley has promised only to explain infrastructure, so he is within his rights to speak of music elements having 'feelings and desires' if they are outside the realm of infrastructure. Yet the book ends with an unfounded pronouncement that given the 'tremendous range of resources available within the common-practice musical system . . . there is nothing fundamentally mysterious or inexplicable about the power of common-practice music' (p. 356). In spite of the ending, in terms of advancing understanding of the cognition of basic musical structures, this book may be the best so far. It is not as good as it gets hopefully, but with the help of experimental psychology, the best so far can get better.

## *General critique: needed – a canon for music cognition*

In its intellectual commitment, focus and productive outcomes, *The Cognition of Basic Musical Structures* is inspirational. It is an example that a young scholar can tackle and partially resolve big problems. Books like Temperley's, written by musicologists and filled with intuitions and speculations about how the mind processes music, provide music cognition laboratories with a resource of hypotheses to last for decades; see also Cohen's (1999) review of Cook (1998).

In spite of concerns that can be raised about the lack of empirical basis for

introspection, much of the value of the book for researchers in music psychology is Temperley's introspection, the detailed account of how he hears and thinks about the structure of music. Whereas Temperley's honest reflection and attempt to model his own hearing may be well deserving of a place on the curriculum of graduate studies in music cognition, his representation of experimental cognitive psychology and its potential contribution to the discussion seems lacking. He portrays the experimental approach as something like an exchange of gifts: one gives a subject a stimulus and the subject gives back a response. 'The methodology of cognitive psychology is primarily experimental: human subjects are given stimuli and asked to perform tasks and give verbal reports, and the psychological processes are inferred from these' (p. 4). Notions of control, data, central tendency, variability, reliability, and statistical analysis are never part of this picture. Musicologist David Butler's (1992) *The Musician's Guide to Perception and Cognition*, in contrast, thoroughly gives experimental methodology its due, beginning with his first chapter.[6]

Weaknesses in *The Cognition of Basic Musical Structures* may in part be the fault of the discipline of music cognition, in particular, its lack of a canon. There is much said about the canon of common-practice music. Is it not time for a canon for the field of music cognition? The history of the discipline of music cognition has yet to be synthesized and recorded, and until it is, new scholars will be presented with an incomplete history, and there will be a continual rediscovery of ideas and an inadequate representation of the past. A more encompassing appreciation of the discipline of music psychology, its foundations and history, might have eliminated some of the problems with the book: a somewhat presumptuous title (implying more cognitive psychology and less computation than is actually delivered), rediscoveries of cognitive wheels, and failures to acknowledge some past work. These weaknesses send a message to the field of the Psychology of Music, or to Music Cognition in particular. The field is characterized by many edited volumes, each of which advances the field, but without synthesis. When a single-authored book appears, one hopes for that synthesis. Whereas Temperley's book title suggests a synthesis, instead what is found is its need. Temperley's book is in one sense much like any other set of edited readings, a collection of his writings from very different sources plus some added essays. Yet it begins to differ from these collections in its four-step methodological approach (identification, introspection, computation and test) to each of the six basic music structures identified. Empirical data however should replace or add to the introspection.

Temperley's insights about music cognition are on the mark, but seasoned cognitive psychologists may twiddle their thumbs during his unreferenced postdoctoral musings: that music is about pattern recognition (early pointed out by Simon and Sumner, 1968); that to build something is to understand it (Miller et al., 1964); that optimal complexity saves us from boredom or chaos; that 'enjoyment of a piece is affected by familiarity to some extent'

(Temperley, p. 235; cf. Berlyne, 1971); that repeated rhythms are easier to recognize than those without repeated elements (Garner, 1974; Royer and Garner, 1970); and that 'surprise occurs because our expectations have been denied' (Temperley, p. 231), a basic tenet of information theory and fundamental to Narmour's (1990) theory. Temperley's musings may be fresh for the younger generation whose members too may be without the history. In short, one missing aspect of the book is the foundation of cognitive psychology that one learns in introductory texts and courses or comes to appreciate over the years. To Temperley, history seems to begin with Lerdahl and Jackendoff (1983) but in his defense it may be argued that if an author now worried about doing real justice to the past of music cognition, nothing would ever get written.

In conclusion, Temperley achieves his goal of advancing the understanding of music through identification, description and modeling of components of musical infrastructure. His book itself builds an infrastructure that supports new bridges between the fields of cognitive psychology, music and computer science. It also provides a foundation for appreciating his next works, which have already started to appear. The book has much merit, and the MIT Press should be commended for adding this work to its stellar catalogue in the area of music cognition.[7]

ACKNOWLEDGEMENTS

Appreciation is extended to Dr Bradley Frankland, Dr Henrietta Lempert and Dr Warren Brodsky (Book Review Editor) for comments on an earlier version of the review.

NOTES

1. Detailed examples of music and algorithms are available on Temperley's website (http://www.link.cs.cmu.edu/cbms/index.html), but the reference to this resource is buried in footnote 10 of chapter 1 (p. 362) and the specific website location information is no longer correct. This asset seemed worthy of more recognition.
2. The importance of distinguishing between misses and false positives was identified by Frankland (1999) in his empirical test of *GTTM* local grouping rules (see also Frankland and Cohen, 1996a, in press). Frankland suggests that in the *GTTM* framework, misses are more serious than false positives. False positives are to be expected because listeners (or, in this case, Shaffrath, the compiler of the Essen database) report at the level of the tactus (intermediate time base in the metric hierarchy). The theory predicts many boundaries below this level.
3. For example, Cohen (2000) has argued that an appropriate depiction of the probe-tone profile uses a circular x-axis representing the cycle of fifths, in contrast to the usual linear chromatic x-axis. The circular format makes greater sense from the point of view of modulation to nearby keys in tonal music.
4. Empirical evidence that Temperley does not mention supports the fact that a key is established after the presentation of a few notes (Cohen, 1991; Frankland and Cohen, 1996b; Krumhansl, 1990).

5. Bartlett (1932) is credited with first proposing the concept of schema to account for results in his studies of memory that showed recall of details of stories that were plausible but not originally present. Implicit in this concept was the difference between memory for surface and meaning of a text. Later definitions of schema include: 'a data structure for representing generic concepts in memory' (Rumelhart, 1980: 34); 'general knowledge structures that aid in comprehending information (e.g. what sequence of steps one should go through while ordering food in a restaurant)' (Sternberg and Ben-Zeev, 2001: 53); 'An organizing tendency that is distilled from past experience with an object or an event. Schemas are used to guide memory recall' (Matlin, 2002: 501); 'A pattern of knowledge describing what is typical or frequent in a particular situation. Thus a "kitchen schema" would stipulate that a stove and refrigerator are likely to be present, whereas a coffeemaker may be or may not be present, and a piano is likely not to be present' (Reisberg, 1997: 620). The tonal hierarchy, the major triad, and an interval could be as readily described by the term as could larger concepts such as a cadence or a musical topic.
6. Indeed it is surprising that Butler's (1992) book is not referred to by Temperley, who spent two years at Ohio State University in the same Department where Butler taught. Butler is acknowledged by Temperley in the preface.
7. In a book of this size and scope, a few editorial concerns can be expected, e.g. 'somwhat' and 'none seem' (p. 12), 'in the exactly same way' (p. 42), 'is to establish is' (p. 59), 'a interesting' (p. 231), 'can convey a create a kind' (p. 340), 'scores . . . could not be scored' (p. 44). Brackets around publication dates are sporadic throughout the Notes section and also for example, p. 30. Colloquial phrases like 'which is no disaster' (p. 98) and 'unimaginably huge number' (p. 99) are awkward in the general scholarly context. A glossary for terms like parallelism, hemiola, meter, -etic, -emic, and rock would be appropriate for a multidisciplinary readership.

## REFERENCES

Bartlett, F.C. (1932) *Remembering: A Study in Experimental and Social Psychology*. London: Cambridge University Press.

Berlyne, D.E. (1971) *Aesthetics and Psychobiology*. New York: Appleton-Century-Crofts.

Butler, D. (1992) *The Musician's Guide to Perception and Cognition*. New York: Schirmer.

Cohen, A.J. (1991) 'Tonality and Perception: Musical Scales Primed by Excerpts from The Well-Tempered Clavier of J.S. Bach', *Psychological Research/Psychologische Forschung* 53: 305–14.

Cohen, A.J. (1999) 'Musicology Alone?', review of Nicholas Cook's *Analysis of Musical Multimedia, Music Perception* 17: 247–60.

Cohen, A.J. (2000) 'Development of Tonality Induction: Plasticity, Exposure, and Training', *Music Perception* 17: 437–59.

Cook, N. (1998) *Analysis of Musical Multimedia*. Oxford: Oxford University Press.

Costa, M., Bitti, P.E.R. and Bonfiglioli, L. (2000) 'Psychological Connotations of Harmonic Musical Intervals', *Psychology of Music* 28: 4–22.

Cuddy, L.L. and Cohen, A.J. (1976) 'Recognition of Transposed Melodic Sequences', *Quarterly Journal of Experimental Psychology* 28: 255–70.

Deutsch, D. and Feroe, J. (1981) 'The Internal Representation of Pitch Sequences in Tonal Music', *Psychological Review* 88: 503–22.

Frankland, B.W. (1999) 'Empirical Tests of Lerdahl and Jackendoff's (1983)

Low-Level Group Preference Rules for the Parsing of Melody', doctoral dissertation, Dalhousie University, 1998, *Dissertation Abstracts International: Section B: The Sciences & Engineering*. 60(3-B) 1323 (UMI No. NQ36580).
Frankland, B.W. and Cohen, A.J. (1996a) 'Boundary Formation in Melody: Effects on Memory', in B. Pennycook and E. Costa-Giomi (eds) *Proceedings of the 4th International Conference on Music Perception and Cognition*, pp. 143–7. Montreal, August.
Frankland, B.W. and Cohen, A.J. (1996b) 'Using the Krumhansl and Schmuckler Key-Finding Algorithm to Quantify the Effects of Tonality in the Interpolated-Tone Pitch-Comparison Task', *Music Perception* 14: 57–83.
Frankland, B.W. and Cohen, A.J. (in press) 'Parsing of Melody: Quantification and Testing of the Local Grouping Rules of Lerdahl and Jackendoff's (1983) *Generative Theory of Tonal Music*', *Music Perception*.
Garner, W. (1974). *The Processing of Information and Structure*. Oxford: Erlbaum.
Halpern, A. (1989) 'Memory for the Absolute Pitch of Familiar Songs', *Memory & Cognition* 17: 572–81.
Koestler, A. (1989[1964]) *The Act of Creation*. London: Arkana.
Kostka, S. and Payne, D. (1995) *Workbook for Tonal Harmony*. New York: McGraw-Hill.
Krumhansl, C. (1990) *The Cognitive Foundations of Musical Pitch*. Oxford: Oxford University Press.
Lerdahl, F. and Jackendoff, R. (1983) *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press.
Levitin, D.J. (1994) 'Absolute Memory for Musical Pitch: Evidence from the Production of Learned Melodies', *Perception & Psychophysics* 56: 414–23.
Marvin, E.W. and Brinkman, A.R. (2000) 'The Effect of Key Color and Timbre on Absolute Pitch Recognition in Musical Contexts', *Music Perception* 18: 111–37.
Matlin, M.W. (2002) *Cognition*. Orlando, FL: Harcourt.
Meredith, D. (2002) 'Review of *The Cognition of Basic Musical Structures* by David Temperley', *Music Scientiae* 6: 284–302.
Miller, G. (1956) 'The Magical Number Seven Plus or Minus Two: Some Limits on Our Capacity for Processing Information', *Psychological Review* 63: 81–97.
Miller, G., Galanter, E. and Pribram, K. (1964) *Plans and the Structure of Behavior*. New York: Holt, Rinehart & Winston.
Moore, A. (1993) *Rock: The Primary Text*. Buckingham: Open University Press.
Narmour, E. (1990). *The Analysis and Cognition of Basic Melodic Structures: The Implication–Realization Model*. Chicago: University of Chicago Press.
Platt, J.R. (1961) 'Beauty: Pattern and Change', in D.W. Fiske and S.R. Maddi (eds) *Functions of Varied Experience*, pp. 402–30. Homewood, IL: Dorsey.
Povel, D.-J. and Essens, P. (1985) 'Perception of Temporal Patterns', *Music Perception* 2: 411–40.
Reisberg, D. (1997) *Cognition*. New York: Norton.
Royer, F.L. and Garner, W.R. (1970) 'Perceptual Organization of Nine-Element Auditory Temporal Patterns', *Perception & Psychophysics* 7: 115–20.
Rumelhart, D.E. (1980) 'Schemata: The Building Blocks of Cognition', in R. Spiro, B. C. Bruce and W.F. Brewer (eds) *Theoretical Issues in Reading Comprehension*. Hillsdale, NJ: Erlbaum.
Schaffrath, H. (1995) 'The Essen Folkson Collection', ed. D. Huron. Stanford, CA: Center for Computer-Assisted Research in the Humanities.
Simon, H.A. and Sumner, R.K. (1968) 'Pattern in Music', in B. Kleinmuntz (ed.) *Formal Representation of Human Judgment*, pp. 219–50. New York: Wiley.

Sternberg, R.J. and Ben-Zeev, T. (2001) *Complex Cognition: The Psychology of Human Thought*. New York: Oxford University Press.
Temperley, D. (1999) 'What's Key for Key? The Krumhansl-Schmuckler Key-Finding Algorithm Reconsidered', *Music Perception* 17: 65–100.
Temperley, D. and Bartlette, C. (2002) 'Parallelism as a Factor in Metrical Analysis', *Music Perception* 20: 117–49.
Terhardt, E. and Seewann, M. (1983) 'Aural Key Identification and its Relationship to Absolute Pitch', *Music Perception* 6: 63–83.

**Annabel J. Cohen**
Department of Psychology
University of Prince Edward Island, Canada

D. COPE, *Virtual Music: Computer Synthesis of Musical Style*. Cambridge, MA: MIT Press, 2001. 345 pp. + 5 Appendices 200 pp. ISBN: 0–26–203283-X (hbk) $50 (£33.50)

Creating compositions in a particular known style is not only interesting to students of composition, but also to psychologists seeking to understand how the mind organizes musical material as well as for Artificial Intelligence (AI) developers wishing to develop automated composition software. Each of these groups can benefit from sharing in practical and theoretical developments. In AI development, several recent volumes on the topic have been published, including two volumes with a neural network focus (Griffith and Todd, 1999; Todd and Loy, 1991). In *Virtual Music* David Cope explicates a most significant contribution to this field. For those already familiar with Cope's software, Experiments in Musical Intelligence (EMI), the current volume offers little in the way of new technical information. Rather, in the book Cope has continued to refine his software and has given generous airing about the program and its philosophical aspects from a variety of specialists. The technical details of the software are sandwiched between chapters focusing on far reaching philosophical implications as well as lay descriptions of its operation. This format stems from the colloquia held about the program and the ideas of computational composition. The book also includes sections of commentary by Douglas Hofstadter, and various perspectives and analyses by Eleanor Selfridge-Field, Bernard Greenberg, Steve Larson, Jonathan Berger, and Daniel Dennett.

The first chapters give a strong indication that the challenge of computer based composition mimicking the style of past composers has almost been met. The chapters consist of a basic description of how Cope's program works, and some more philosophical perspectives. The opening chapter describes a game which is a direct test of the software which Cope developed. The listener is asked to determine which of four pieces are composed by a

target composer and genre (Bach Chorales in one game, and Chopin Mazurkas in another) and which by the software. For those who lack the musical score reading and performing skills to play the game, a CD is included containing numerous examples in addition to the games.

Hofstadter commences the second chapter by admitting that some of the derisive predictions he had to make about computers being able to compose in *Gödel, Escher, Bach* (1979) had been proven wrong. For example, he suggested earlier that computers could never outplay great chess players. In discovering that computers can outplay humans at chess Hofstadter philosophizes that the process of playing chess 'is more mechanical than we had realized' (p. 36) but laments that computational mimicry of beauty in music is next in line. Indeed the tone throughout the chapter is a search or hope for some depth of humanity which computers can never attain (with a response to this notion in the following chapter by Cope). Hofstadter was so impressed by Cope's program that he used it (and the game) in his lectures. It led him to consider the issues which emerge from the architecture of automated composition emulation – in particular the principles of how the program worked, and what this meant for music and human cognition. One of the numerous subsection headings summarizes the argument: 'Is music Just Splicing of Licks, and No More?' (p. 41). At a simple level the answer is: Yes! The software has as its input a selection of works by a target composer and a target style and genre. From this, the software breaks the music into components – which Hofstadter refers to as 'chopping up'. The chopped up portions are then reassembled. The process is more complex, but this gives the general outline of the program's operation. Hofstadter focuses on the reassembly parts of the program. The principles of how to reassemble the various pieces to create a new work from the source material are described in a very untechnical manner. The musical insights Cope required to understand how this process is programmed, and therefore how the process can be understood 'may in fact constitute a new contribution to music theory' (p. 46) according to Hofstadter. He describes how the complexity of EMI is gradually increasing and becoming better at emulating target composers. With this, he returns to his concerns about the ability of a computer to communicate emotional expression (pp. 54–5). In support of EMI as a model of mind, he suggests there exists modular brain specialization in composition. Later in the chapter, he also discusses the dilemma of AI replacing the human brain – that music is simply pattern, reducible to a purely mechanical process (p. 80).

One of the threads running through the book is the convincingness of the EMI compositions. Hofstadter recounts an email from a member of his audience (a composer) who remained unconvinced with Bach examples of the 'name the composer game,' but was quite amazed by the Chopin compositions. The comforting thing about the presence of the Hofstadter chapter and those of invited guests is that the findings which Cope mentions regarding the ability of EMI to fool its audience is tempered but also given a sense

of increased validity. Hofstadter discusses how his own compositions were interpreted by EMI to produce a new piece in the style of Hofstadter. In the view of the composer the project had limited success. But this merely supported Hofstadter's own model of the program that, for one very familiar with a compositional style, EMI still has some work to do. In Chapter 15, Dennett describes another way of modeling this meta-understanding, through the detection of the 'muselot' fallacy, the musical counterpart of 'gramelot' – where spoken words appear as a foreign language to all but those who speak that language (pp. 290–291). So for Hofstadter, although the EMI composition is still akin to muselot, there is no longer a sense that passing this Turing test of musical-composition-authenticity is unobtainable, as there was in *Gödel, Escher, Bach*.

Chapters 4–6 present the nuts and bolts of EMI, with worked examples of the principles upon which the program operates. EMI provides insightful and provocative solutions to how the mind stores musical information. If Cope (perhaps with an assistant) expanded the inner chapters (especially chapters 4–6) by providing a detailed harmonic, melodic and structural analysis to the many musical examples, I think we could see an elegant alternative to understanding problems which plague undergraduate music students who study harmony. By seeing music as patterns and understanding the regularities and idiosyncrasies among them, I suspect struggling music students might become enlightened about the purpose of harmonic analysis – at least this has been my own experience as a teacher of undergraduate harmony. To produce software which mimics musical style requires deep insight into how music is structured (perhaps with the exception of self-learning algorithms, but even then), and so a bridge from the software engineer to the music theorist could provide an alternative and beneficial pathway to such understanding.

Chapter 4 commences with the mechanics of the principles upon which the software works. It provides a demonstration of how EMI can break up Bach chorales into small units, in this case at the beat level, and then recombine these units to produce a new Bach chorale. The program is able to maintain structural logic in the recombination process by storing large databases of information about suitable progressions for each of the musical units. The software selects one of the Bach chorales as its model to assist with the structural layout. That is, the form (phrasing and sections) will resemble this template, but the beat-to-beat material will come from any of the chorales. This structural logic is a watershed of apparent authenticity. It mitigates the possibility that the newly composed piece will sound like a meandering, directionless pastiche of musical ideas. In the chorale composition example the structure at a global level is maintained by copying the cadence progressions at each phrase (summarized nicely in Figure 4.14 on p. 103). To enhance the probability of successful recombination, the software also has in its database variations, or 'transformations' of the stored units.

One of these is the diatonic transposition of the original music. For each stored musical unit this produces an additional six options per unit (one for each scales degree). It allows the program to achieve greater structural unity and authenticity because it is not fully dependent on producing just patchworks of source material. As discussed later in the book, an important application of this kind of transformation is in producing musical sequence. That is, these transformations may be viewed as a direct application of compositional devices to support the recombination process.

Chapter 5 discussed an even higher level of sophistication of EMI which endeavors to ensure that the musical style of the target composer and genre are maintained. The software is able to identify patterns which occur across compositions by a single composer and patterns which occur across historical periods in music. Cope argues that there are around four to ten occurrences (including variations) of these signatures per movement of music, and that a signature is usually two to five beats long. The software searches for signatures by pattern matching algorithms and ensures that they will occur in the newly composed work. The chapter presents the various signatures found in Mozart piano sonatas, for example the $I^6_4$-*V*-*I* progression. Such a harmonic signature is labeled a 'harmonic underpinning.' Other forms of signatures in this example include particular rising scales, resolving seconds and falling leaps, as identified by the pattern matching process. Another aspect connected with the identification of style is what Cope refers to as an 'earmark.' An earmark is a musical convention which is usually unique within the movement. It can serve to mark global position within a movement, and help to identify important structural points in the music. An example in the Mozart piano concertos is the change in character and harmonic preparation during the transition to the cadenza. Such patterns, according to Cope, have little musical interest (as distinct from signatures) but provide global, structural markers. These markers need to be retained in the newly composed music at structurally required points in the piece. Texture and harmony are more important than the actual notes used, making earmarks 'principles rather than data' (p. 121). While Cope discusses the problems in applying earmarks in EMI, he claims that more recent versions of the software have improved the problem, but no further details are provided (p. 121). Throughout the book there is an underlying message about the reducibility of music. At the end of Chapter 5 we see the most forceful articulation of this message in the section on unification, that all music is pattern (p. 125), a point which caused Hofstadter some concern.

Chapter 6 describes further sophisticated elements of the software. This includes the description of algorithms for detecting musical boundaries, dealing with the hierarchical levels of the piece, and the Schenkerian style structural analysis which describes musical materials as consisting of groups reducible to their tension relationships – a statement, preparation, extension, antecedent or consequent (referred to by the acronym SPEAC). This process

allows calculation of tension and enables EMI to make decisions about the functional relationship within a musical structure, thus supporting more convincing selections of material – that is, material which maintains an 'authentic' tension interrelationship. The SPEAC labeling of a fragment of music does possess a degree of subjectivity (p. 133). Perhaps this could be addressed in future versions by using the quantified tension calculations, or considering other established methods, such as that by Lerdahl (1996, 2001). Cope concludes this chapter by acknowledging that the complexity of the software may be overwhelming. As Hofstadter's concentric circle model proclaims, the closer we get to the ultimate imitation, the greater the complexity.

The second section of the book (chapters 7–10) further tunes the microscope into the details of operation of EMI. In Chapter 7 the coding used for recording and manipulating musical events (they resemble MIDI) and the guidelines for choosing pieces for the database are discussed. In Chapter 8 Cope explains how the various patterns are normalized so that they can be smoothly combined. For example, if a pattern in one segment in one piece is a rhythmic augmentation of a segment in another piece, the rhythm is adjusted so that the two are compatible. Cope calls this normalization process 'clarifying' the data. Chapter 9 has a detailed, worked example, showing the bar-by-bar evolution of a 'new' Mozart piano sonata second movement. The example integrates the previously described processes.

Chapter 10 gives a very brief description of the interface – perhaps too brief to be of any great benefit to one who wishes to use the software. It leaves one asking many questions about the specifics of the interface, and if one does not need to know them, perhaps then the chapter could have been omitted altogether. While Cope's intention is just to give a feel for the EMI program, it may just be a little too tantalizing! The chapter could have benefited from a worked example of how to set up source material and then generate a new composition. A reference for more information about the software (such as an internet website) is required here. Toward the end of the chapter the importance of using a human performer is discussed, instead of using deadpan MIDI playback. Cope intends to add expression in future versions. Perhaps he will capitalize on the work already available in computationally modeling musical expression (Juslin et al., 2002). The chapter concludes with the discussion of an example where EMI has failed to produce convincing output of Mozart. It is an example of Cope's self-critical approach, with the 'failed' music presented in the appendix and on the CD.

The third section of the book consists of seven chapters by 'guest speakers' on a selection of topics related to the implications of EMI, followed by Cope's response chapter. Chapter 11, by historical musicologist Eleanor Selfridge-Field, describes a time line showing the development of various musical models of composition from the celestial through to the cognitive model. She then discusses Kivy with respect to emotion in music before elucidating upon

historically relevant mathematical and linguistic prototypes. For example, she explains the eighteenth century musical dice game presented by Kirnberger where after a throw, a table directs the 'composer' to generate a melody in the style of a Polonaise or Minuet. While Selfridge-Field suggests that this game might be a precursor of EMI, the sophistication of Cope's program can also be traced from more complex games, which Selfridge-Field explains. Even a biological model is discussed which might be applicable to EMI, raising issues of 'genetic' recombination and cloning in music.

A current problem with EMI is that the choice of keys is not automated. This can produce several problems, among them, fingering. The example Selfridge-Field uses could be a little clearer. The musical example (Figure 11.6) indicates a fingering problem for a piece generated in the style of a Bach invention. It seems to me that there is a reasonable (though perhaps non-idiomatic) solution at this point as it stands. Either more explanation was required, such as a statement about the fingering alternatives, or a better example could have been chosen.

One of the recurring themes of the book, though not always an explicit one, is how EMI can and does inform us about the compositional process. Selfridge-Field argues that 'existing explanations tend to be both prescriptive and proscriptive (Do this . . . don't do that)' (p. 206). She suggests that the composer's ideas which cannot be explained in such a way can be filled in by a 'black box of imagination.' Combinatorics, of the kind used in EMI and Berggren's approach (introduced by Selfridge-Field), may provide an alternative perspective to understanding the process of composition. She also uses EMI's variety of styles to strengthen the need for looking at the AI literature to better understand the compositional process, and its subsequent analysis because it may lead to the examination of unusual or traditionally poorly explained aspects of a composer's idioms. In the case of Joplin 'it can be seen that tertiary dominants – the dominant of the secondary dominant – could occur. Descriptions are not easily found in harmony textbooks, though, because Joplin does not belong to any recognized canon of distinguished harmonizers. Thus, fidelity to actual repertories can expand our working store of theoretical knowledge, where formal music theory has failed to do so' (p. 210).

The diversity of styles which EMI can emulate is also discussed, including Broadway show tunes, where even the text is simulated. Also discussed is the performer – artificial performance by a literal MIDI player, for example, tends to sound more like a computer composition than if the same work were performed by a non-metronomic human. Towards the close of the chapter, Selfridge-Field suggests that EMI is not directly creative (p. 216) – a point which is revisited in some detail by Cope in the final section of the book. Selfridge-Field presents a model of composition activity which will be of use to those studying compositional processes. An interesting aspect of the model is that it draws on the dichotomy of a library of stylistic traits (memory or databases of previous pieces, if you like) and the procedural concepts for

creating (knowledge of rules related to form, structure and other aspects of composition). EMI is based more strongly on the former (memory), and Selfridge-Field suggests that for humans there may be too much focus on the latter (procedures).

Chapter 12 is by Bernard Greenberg, a software engineer and lover of Bach. He briefly introduces the issue of assessment which could have been looked at further, given the broad scope of analysis Cope's software undergoes in this book. That is, as well as investigating how and why EMI seems to work so well, how are we assessing that it is doing these things so well? Only lightly touched upon in some parts of the book are issues related to cultural values and historical contexts. But the problem is important because if certain non-EMI related parameters are manipulated (such as knowledge of the composer versus anonymous, as in the game), the assessment of the compositions can be drastically changed (see also my comments about the last chapter, at the end of this review). I would have liked to see more discussion on how culture and personal backgrounds and belief systems influence the various kinds of responses to the pieces. Greenberg intends to present in this chapter an examination and better understanding of the software, Bach's music, and the compositional process – 'three intimately conjoined domains' (p. 221). He mentions the software's ability to fool the audience and hence pass the Turing test. Again the modeling of emotion is discussed, but significant recent advances might need to wait for a second edition. Greenberg's passion for Bach goes a little overboard if not patronizing at times to the point of digression, implying that it helps to be a Christian (see also p. 234) to understand the full breadth of Bach's *Matthäuspassion* or that 'Wagner is no Bach.' He spends several pages discussing the limitations of EMI in dealing with counterpoint. He then provides a lovely example of the dissonances which can arise from counterpoint, referring to a bar from BWV 787 and BWV 622. In sum, Greenberg identifies flaws in Cope's program due to its inability to deal with the linear processes of counterpoint. This apparent weakness in EMI makes me think not so much of its limitations, but more the amazingly detailed knowledge some people have of the compositions of a particular composer. The concentric circle model described earlier in the book by Hofstadter seems apt to describe the situation. Greenberg is probably closer to the centre of the circle than is EMI.

In Chapter 13 Steven Larson, a music theory teacher at the University of Oregon, provides a detailed report to a student about an assignment in which the student was to write fifteen two-part inventions in the style of Bach. The student was EMI. The chapter once again invokes the idea that the more attuned and familiar the listener (in this case, Larson) to the works of the composer being emulated, the easier it is to detect the 'plagiarisms' which EMI tries to hide. And again, this does not demonstrate a negative aspect or failure by EMI, but rather a reflection or model of the possible processes humans may use in composition. For example, Larson acknowledges that

'Bach recycled his own material, too' (p. 238). However, Larson is also genuinely impressed by the way EMI has absorbed and applied the essential principles of Bachian two-part invention writing. This is a highly educational chapter, one worth inspection by students learning first species counterpoint. Larson even includes a recipe for a Bach invention (p. 249). In essence, he attempts to un-reverse the reverse engineering EMI uses to compose two-part inventions, thus providing a traditional analysis of the EMI output.

Among the commentators of Cope's book is Jonathan Berger, a music theorist at Stanford University. In Chapter 14 he discusses creativity and touches on the problems of subjectivity and expectation, and how these could be computationally modeled. Berger devises a list of symbols used to represent different classes of expectation, suitable for annotating different conditions of expectation and realization on a musical score. Later in the chapter squares representing amount of expectation (or activation) are used, with larger squares showing stronger expectation of a particular pitch. Berger claims to have developed a measure of the disparity between expectation and realization which he refers to as 'degree of realized expectation.' The latter notation system stems from the focus of this chapter being on a neural network, developed by Berger and Dan Gang. The network attempts to model the listening process and is called 'Experiments in Musical Listening.' Given the central nature of neural networks in this chapter and the computational modeling theme of the book, the novice reader might appreciate some further background to neural networks, as is the case in the mention of neural networks by Cope in Chapter 1 (p. 11) where only the back propagation architecture is cited (e.g. Gjerdingen, 1990; Todd and Loy, 1991). Berger does pay tribute to the pivotal contribution by Catherine Stevens in which a network learns the *Blue Danube* melody. And he presents an interesting augmentation of Stevens' design by introducing nodes representing the beats of a meter. Of course it is doubtful that such a network can represent sufficient data for modeling the listener, but if Berger intends to operate like Cope, this is probably best interpreted as work in progress with many more amendments required to model the hierarchical, emergent and veridical factors which are likely to be in operation during the music listening process.

Incorporating the contribution of rhythm in expectation is the single most important addition made in Berger's model with representation of 'harmonic progressions but this time with metric and beat position represented' (p. 276). The example demonstrating how the network determines the expectations generated by the hypothetical listener could be made a little clearer. There appear to be some inaccuracies in the harmonic analysis of the EMI generated piece in the style of a Bach Invention, and the meaning of the numbers in 'activation-square' (my expression) notation is not made explicit (does 0 represent C or does it represent G when the piece changes to the key of G?). Also, in the fourth measure of Figure 14.6 (incorrectly referred to as 14.5 on p. 275) the harmonic rhythm of the common-meter piece, to me,

moves in even quavers (eighth notes) rather than the 'syncopated' patterns shown, and the harmonic analysis seems slightly, though significantly (because the expectation outputs shown at the top of the score are dependent on previous harmony), erroneous. Instead of a progression of: G: V-I-(I)-V-I-IV-(IV)-ii, I hear more evenly flowing harmonic rhythm of: G: $V^6$-I-(I)-$V^4_3$-$I^6$-IV-ii-$I^6$ (parentheticals indicating continuation of the harmony across the eighth-note pulse). So, while Chapter 14 had some ground breaking attributes, it seems in need of a little more editing.

Chapter 15 is written by Daniel Dennett, the director of the Centre for Cognitive Studies at Tufts University. He compares EMI with evolutionary genetics and genius, and argues that a genius does not occur in isolation from history and society. I was particularly entertained by his recipe for a St Matthew Passion (p. 285): 'First, make a Bach, and educate it, instilling all the best products of the contemporary culture. Then sit back . . .' (p. 285). Dennett discusses the distinctions between creativity in the virtual world as compared to the real world. In the virtual world the spontaneous and extraneous events which seem to be related to real world creativity must be added by the programmer – what Hofstadter refers to as 'spontaneous intrusions.' Dennet paraphrases spontaneous intrusion as: 'when you're modeling creativity, there should be junk lying around that your creative processes can bump into, noises that your creative processes can't help overhearing' (p. 287). This suggests a clearly random or stochastic component to creativity. But perhaps this looseness is the case rather because we don't understand the process at that level of complexity. It is therefore surprising that more space is not devoted to system dynamic and chaotic models which provide a compact alternative to some aspects of the problem of inducing creativity through AI.

Hofstadter's second chapter in the volume (Chapter 16) is structured as a series of questions and answers that philosophize further on issues of creativity and AI. Throughout this chapter Hofstadter relies on comparisons with language and literature. In the early part of the chapter he does not seem to acknowledge that music is in some important ways different to verbal languages (though amends are made in the answer to the second last question on p. 304). Such languages have more strict syntactic and semantic structures. Music has greater freedom because it does not need to satisfy denotative criteria. So his doomed attempts to generate successful translations of Pushkin's novel in verse, *Eugene Onegin*, into English using sophisticated AI technology seemed like an unfair comparison. With a background in psychology, I found Hofstadter's near outrage about music existing in the mind of the listener old fashioned if not disturbing (p. 301): 'the meaning of a piece of music (or literature) is not invented out of nothing by us receivers; it resides in the mysterious catalytic power of *the sequence of notes* that somehow that composer was able to find, and which other people had never stumbled across before' (p. 302, italics in original). It is a shame that the possibility of cultural learning is not taken as an additional constraint about what gives music its

meaning. Foucault was mentioned in the previous chapter, and perhaps a visit to this chapter (via his views about enculturation) may have ironed out this view that implied greatness as a magical gift from above (Cope does this himself on p. 312 when discussing Mozart). In short, I was disappointed by this chapter given what a terrific thinker Hofstadter is.

Cope provides his own review in Chapter 17, making comments about selected portions of the third section of the book. From a music-psychology perspective, he makes one of his most important assertions about EMI suggesting that 'to some extent all composers use recombinacy as a compositional process' (p. 309). As such, Cope has developed an AI approach which has led to the emergence of a psychological theory of how music is stored and combined during the composition process. This much needed theoretical proposal (outlined in the book and in Cope, 1996) presents important research fodder for future investigations on composition and creativity. An example of how this provides a theory of creativity and composition comes from one of Cope's own experiments, where he fed an EMI-generated work back into EMI. By continuing this process, Cope believed that EMI was no longer emulating a composer, but evolving its own, individual style. Of course, there is much to explore here – whether the end product will lead to creation or recursion – but it does provide importantly needed theory. I certainly align with Cope's view that a machine can be creative. The narrower alternative is a 'singularly homocentric and seriously limited view of creativity' (p. 310).

In his response to Greenberg's Chapter 12, Cope discusses issues of musical worth. He argues that popularity and musical worth are not the same thing, providing the 'banal and repetitive melody' (p. 314) of *In the Mood* performed by the Glenn Miller Band as an example. The success of this piece derives from 'blind luck' (p. 314) rather than musical worth, according to Cope. His point here is that issues outside the music itself influence one's perception of the piece, what Leonard Meyer calls referentialism (1956). So, in the same light, knowing that a piece of music is composed by a machine can negatively influence the listener's perception of the worth of the piece. Cope develops a gradually more defensive tone in response to issues raised in previous chapters. For me, this tone becomes troubling during his response to Hofstadter's Chapter 16, specifically the issue of music happening in the audience's mind. They both attack this notion with a somewhat romantic view of music, that there has to be more to it – something outside the mind but within the music perhaps; that music happening in the mind does not explain how it gets its meaning. It is something more than the composer's intention, according to Cope, perhaps something mysterious or mystical (he implies, but does not state). However, what is not considered here, as I have mentioned earlier, is the role of historical context and culture in creating musical meaning – a consideration riddled with complications, but not in exclusion of music existing in the mind because these contextual and cultural cues must also be represented in the mind.

The most refreshing aspect of this book is the recurring theme which is captured in the opening of the final Chapter 18: 'Since 1987 . . . the program continues to challenge many listeners' assumptions about creativity, inspiration, and how and why we listen to music' (p. 323). This explains his earlier mention of how his software delights and angers people (p. 139). Much of this chapter focuses on creativity and provides important resources for psychologists studying the phenomenon. Cope presents a convincing argument that supports the view that EMI *is* creating music, rather than just recombining it. A section headed 'Music Cognition' discusses Dowling and Harwood's (1986) post-Peirce classification system of index, icon and symbol, and how this classification fits in with EMI. The section also explains how and why different listeners might react to EMI compositions. Also in this section Cope explains how the seeds for the architecture of EMI were planted in his mind by the segmentation into 'a series of small episodes' (p. 333) he experienced as a child through listening to 78 rpm discs of Rachmaninov's Second Piano Concerto. He expands his argument about creativity by citing important examples both in music and research literature where copying or borrowing is a natural part of the creative process. The last half dozen or so pages of the chapter see the return of some slightly bitter tones. It gave the impression that Cope has had to work hard at defending his ideas and software. But I was convinced, quite early on in the piece, of the worth of the software and the reworked perspective I now have about creativity, compositional processes and a refreshing view of how we might fragment, store and recombine musical information.

In sum, this book has an interesting, unusual structure, mixing lay-discussion and philosophy with technical information and detailed musical examples. A background in music composition, philosophy and AI would be necessary to gain the most out of this book. Yet, while other publications about EMI cater for more specialized audiences (Cope, 1991, 1996), this volume contains something meaningful for people with any of these backgrounds. There are almost copious musical examples, some which could be explained in a little more detail, but the inclusion of numerous examples in the appendices and the accompanying audio CD is most generous. Unfortunately, Appendix D is very long, and requires a more detailed listing of its content; I spent too long thumbing through this section looking for pieces cited in the chapters.

Importantly, by discussing the process through which the software produces its composition and the philosophical issues, *Virtual Music* presents music psychologists with some important and fascinating frameworks in dealing with cognitive models of creativity and musical memory, something that will benefit empirical research. Therefore, while this book is not intended as a psychological text, it is worth examination for researchers in these fields. In content and presentation it is a fascinating and provocative volume.

REFERENCES

Cope, D. (1991) *Computers and Musical Style*. Madison: A-R Editions.
Cope, D. (1996) *Experiments in Musical Intelligence*. Madison: A-R Editions.
Dowling, W.J. and Harwood, D.L. (1986) *Music Cognition*. New York: Academic Press.
Gjerdingen, R.O. (1990) 'Categorization of Musical Patterns by Self-organizing Neuronlike Networks', *Music Perception* 7: 339–69.
Griffith, N. and Todd, P. (eds) (1999) *Musical Networks: Parallel Distributed Perception and Performance.* Cambridge, MA: MIT Press.
Hofstadter, D.R. (1979) *Gödel, Escher, Bach: An Eternal Golden Braid*. Hassocks: Harvester Press.
Juslin, P.N., Friberg, A. and Bresin, R. (2002) 'Toward a Computational Model of Expression in Music Performance: The GERM Model', *Musicae Scientiae. Special Issue 2001–2002*: 63–122.
Lerdahl, F. (1996) Calculating Tonal Tension', *Music Perception* 13: 319–63.
Lerdahl, F. (2001) *Tonal Pitch Space.* New York: Oxford University Press.
Meyer, L. B. (1956) *Emotion and Meaning in Music*. Chicago, IL: University of Chicago Press.
Todd, P.M. and Loy, D.G. (eds) (1991) *Music and Connectionism.* Cambridge, MA: MIT Press.

**Emery Schubert**
School of Music and Music Education
University of New South Wales, Australia