Chapter 16

Review of Assumptions

It is likely that reading this book has challenged some of the assumptions that you might have made about music and what music is. My theory not only challenges common sense assumptions, it also challenges many assumptions made by music scientists, music philosophers and music theorists in their efforts to understand what music is.

The theory contains its own assumptions. They do not necessarily all stand or fall together. So in this chapter I list them individually, discuss the consequences of each one, and where relevant list some of the alternatives.

16.1 General Assumptions

16.1.1 Information Processing

Assumption: The human brain is an information processing system. Functional components of an information processing system are also information processing systems. An information processing system can be characterised in terms of inputs, outputs, calculation and storage of information. Processing of information can and should be described in terms of the meaning of information.

Unless one is disposed towards mystical understandings of the human mind (and even perhaps if one is so disposed), there is not much reason to doubt that the human brain is an information processing system. But discussions of the deeper mysteries of human existence sometimes lose sight of this fact.

Music is not an information processing system. Music is the information. Those parts of the human brain that respond to music are the information processing system. So we can ask:

- What is the input? (The obvious answer is music, but we might ask if that is the only input, or if it is the intended input.)
- What is the output? (Emotions?)
- What calculation is being performed?
- What information (if any) is stored by the system?

Some theories of music suppose that the importance of music has to do with its creation. In which case the input/output questions are the other way around:

- What is the input? (Emotions?)
- What is the output? (Music?) What does the output mean?

16.1.2 The Importance of Musicality

Assumption: The perception of musicality is the major result of the perception of music.

For anyone who studies music scientifically with commercial motivations, the importance of musicality is obvious. If you can discover an algorithm that composes music that a large number of people are going to like a lot, then you are going to get rich.

Existing music science does not completely ignore the issue of musicality, but there is a tendency for it to slip below the radar. One reason for this has to do with universality and political correctness. If you ask why some music is better than other music, a common response is that different people have different ideas about what is good, and it wouldn't be fair to pick a criterion of goodness for one person, and the notion of goodness is so variable and culturally determined that it is impossible to measure, so you might as well forget about it.

The difference between what is moderately good and what is really good is also the difference between the rules of music and the mysterious elements of genius and inspiration—which are assumed to be beyond the reach of scientific investigation.

But to deny or ignore the existence of musicality for these reasons is to give up too easily. If music isn't musical enough, then in practice we don't bother to listen to it, and performers don't bother to perform it. Unmusical music isn't really music, so if we don't understand what "musical" is, then we don't understand what music is either.

The notion of musicality implies one-dimensionality: that it is a number which can be defined and measured on a simple linear scale. We know that the emotional effect of music is not entirely one dimensional—some music has a happy feeling, and some has a sad feeling. It is hard to determine exactly how many different emotions have to be enumerated to account for all the different effects that music can have. The concept of musicality does not deny these extra dimensions, but it does relegate them to a secondary role in the effort to understand what music is and what function it serves.

16.1.3 We Need to Explain Perception of Musicality

Assumption: The biological problem of explaining the purpose of music can be reduced to the problem of explaining the purpose of our ability to perceive musicality.

This has a strong negative corollary: that music in itself *does not necessarily serve any biological purpose*. A large portion of the literature of music science, philosophy and theory assumes one or more of the following:

- That listening to music serves some useful purpose
- That performing music serves some useful purpose
- That composing music serves some useful purpose

Some authors discuss purpose without properly analysing it within the framework of theoretical biology, i.e. having more grandchildren, as discussed in Chapter 2. Among those who stick to a Darwinian framework, there is an endless variety of creative ideas about how listening to, performing and composing music can serve one purpose or another.

My theory says that all of these ideas are both wrong and irrelevant. The only reason that composers compose music is so that it can be performed and listened to. The only reason that performers perform music is so that they and their audience can listen to it. The only reason that we listen to music is because we perceive the musicality of music, and this perception of musicality makes us feel good.

We listen to music because we perceive its musicality, but the purpose of our ability to perceive musicality does not have to be the perception of the musicality of music.

16.1.4 Musicality of Speech

Assumption: Musicality is an attribute of the speech of a person speaking. This attribute is perceived by the listener.

If the purpose of perceiving musicality is not to perceive the musicality of music, then it has to be the perception of the musicality of something else.

The only thing that is similar to music other than music itself (and poetry, as discussed earlier) is *speech*.

Assuming that musicality is an attribute of speech does not immediately tell us what the meaning of musicality is, but it does imply that it is something important for the listener to know. And we can go a long way without necessarily knowing what that important something is.

How similar are music and speech? There is a sense in which most music *is* speech. Turn on your radio, and you will hear songs. Songs are music whose major component is an individual speaker speaking to you the listener. The speech (normally called **lyrics**) doesn't always make a lot of sense, and its melodies and rhythms contain unnatural regularities, but it is identifiable as speech.

Even looking at the differences between speech and music, we can see analogies between the relevant aspects:

- Musical rhythm has regularities that are not found in normal speech rhythm, but rhythm is still an important perceived aspect of speech.
- Musical melody has regularities that are not found in normal speech melody, but melody is still an important perceived aspect of speech.
- Music contains instruments other than the human voice, but where those instruments produce pitch values, the timbres of the instruments have characteristics analogous to the human voice, and in particular to human vowel sounds, because they have harmonic components whose frequencies are integral multiples of the fundamental frequency of the sound. The sounds of percussive instruments (which either don't have harmonic components, or they have harmonics that are not integral multiples of the fundamental frequency) may in some cases be considered analogous to consonants.
- The musicality of music causes an emotional response in the listener. The content of speech can also cause an emotional response in the listener. Where music is song, the emotional effect of the musicality interacts with the emotional response to the content of the lyrics.

16.1.5 Music is a Super-Stimulus

Music is a super-stimulus for the perception of musicality.

Musicality is a perceived aspect of speech, and music is highly contrived "speech", which is contrived in such a manner as to maximise its musicality, and which may also lack those features of speech (such as coherent plausible semantic content) which are not relevant to the determination of musicality.

One consequence of this contrivance is the appearance of the observed regularities of rhythm and melody that occur in music but not in normal speech.

16.1.6 Emotions

Assumption: The perception of musicality is an attempt to determine the emotions of the speaker.

The major effects of listening to music are emotion and pleasure. If we already assume that musicality is a property of speech, then it seems plausible that this perceived property tells us something about the emotions of the speaker.

There is, however, an alternative to this assumption …

16.1.7 Our Emotions, Not the Speaker's

Assumption: The perception of musicality affects our emotional response to the content of the speech. It is a measurement of the speaker's internal mental state, but the emotions that we feel in response to music are *not* an estimate of the speaker's emotions.

This is a rather subtle point. The idea is that something about the speaker's internal mental state *justifies* or *confirms* the emotions that we (as listeners) feel in response to the content of the speech. The most likely candidate for the relevant aspect of the speaker's internal mental state is *conscious arousal*—if the content of speech has some emotional significance for us (and assuming that the speaker is aware of this), then most likely the speaker will be consciously aroused to some degree, even if their emotional response to the content of the speech may be different to ours.

16.1.8 Musicality is Not Emotion-Specific

Assumption: The musicality of music does not determine the specific emotion contained in the listener's emotional response to music.

This follows from the assumption that musicality measures conscious arousal, as conscious arousal is an essentially one-dimensional quantity. If an item of music does seem to specify a particular emotion, the assumption is that something else other than musicality is specifying the particular emotion. For example, it is well known that major chords tend to sound happy and minor chords tend to sound sad. A theory of non-emotion-specific musicality does not have anything to say on the causes of this association—it treats it as a separate problem from the problem of explaining what musicality is. The most that can be said, is that if the characteristics of a tune cause it to express a particular emotion, such as sadness, and the tune has a high level of musicality, then the result will be a strong level of that particular emotion (sadness) felt by the listener.

16.1.9 Musical Cortical Maps

Assumption: Cortical maps that respond to music do not exist for the purpose of perceiving music. They exist for the purpose of perceiving speech.

This assumption is implied by the assumption that musicality is a perceived attribute of speech. Those cortical maps that provide input to the perception of musicality must actually be cortical maps that process and perceive the sounds of speech.

There are various corollaries to this assumption:

- The cortical maps that respond to music are more hard-wired than soft-wired, since they are designed to respond to speech. The importance of speech suggests that many cortical maps would have evolved to serve specific and well-defined purposes in the perception of speech.
- For each cortical map that can be identified as perceiving a particular aspect of speech, the purpose of that cortical map is not to perceive that aspect of music.
- At least some of the cortical maps that respond to music may have a primary function in the perception of speech *other than* the perception of musicality. (However, the **split map** theory weakens this assumption: some of the maps may be evolved copies of other cortical maps. In each case, the original cortical map serves a purpose unrelated to the perception of musicality; the evolved copy remains analogous to the original cortical map, but it has evolved to optimise the perception of musicality.)

The implications of this assumption are most radical when applied to those musical aspects which do not exist in speech, or which are quite different to the analogous aspects of speech.

The most significant of these implications are:

- The cortical map that responds to musical scales has a purpose which is not the perception of musical scales because *there are no musical scales in speech.*
- The cortical map that responds to harmony and chords has a purpose which is not the perception of different notes played simultaneously because *there is only ever one current pitch value in the speech of a single speaker.* (And it seems highly implausible that the human brain has evolved a specific capability for perceiving relationships between the speech sounds of simultaneous speakers.)
- The cortical map that defines and perceives the "home" quality of home notes and home chords does not have the purpose of perceiving home

chords, because there are no chords (and therefore no home chords) in speech. Also, the home chord for a melody normally belongs to the set of preferred home chords for the scale that the melody exists in, and (as already noted) there are no scales in normal speech.

- The cortical map that responds to regular beat patterns has a purpose, and this purpose is not the perception of regular beat patterns because *there are no regular beat patterns in speech.* There are *irregular* beat patterns in speech rhythm, and it is likely that the purpose of the cortical map that responds to regular beat patterns is actually to respond to irregular beat patterns.

- If there is a cortical map that responds to the occurrence of exact **non-free** repetition in music, then the purpose of that map cannot be to respond to exact non-free repetition, since such repetition does not occur in normal speech (with the possible exception of **reduplication**, which consists of non-free repetition of components within a word). The purpose of the cortical map that responds to non-free repetition in music may be related to the occurrence of *approximate* repetitions within speech intonation patterns.

- The cortical map which responds to (or is affected by) rhymes at the ends of lines must have some purpose other than the perception of rhymes at the ends of lines because *normal speech does not rhyme.*

- The cortical map that responds to multiple dancers dancing in synchrony must have some purpose other than the perception of multiple dancers dancing, because *normal speech is only spoken by a single speaker.* (Here we are assuming that the perception of dance is an aspect of musicality perception which has to be included in the theory as an aspect of speech perception.)

16.1.10 Symmetries

There are various symmetries in the perception of music and speech, where a perceived aspect is invariant under some set of transformations applied to the input information.

For each symmetry we can ask the following questions:

- What does the symmetry apply to?

- Does the symmetry serve a **functional** perceptual requirement of invariance, i.e. is the perception required to be invariant under the symmetry, or, is the symmetry an **implementation** symmetry—a consequence of internal mechanisms underlying the perception?

- Is the symmetry a symmetry of both music and speech perception?

Each of the six observed symmetries is defined by a set of transformations:

- Time Translation: Music played later or earlier.
- Time Scaling: Music played faster or slower.
- Pitch Translation: Music transposed into a different key.
- Amplitude Scaling: Music played louder or softer.
- Octave Translation: Adding or subtracting octaves to individual chords and notes within chords (including bass notes). Most musical scales are invariant under translation by an octave.
- Pitch Reflection: Reflecting notes in a scale about a pivot note.

Each symmetry consists of invariance of some aspect of the musical quality of music when the music is transformed by any member of the corresponding set of transformations.

Time translation invariance is seemingly trivial in the sense that the response to an item of music is fairly independent of the time it is played at. A slightly less trivial aspect of time translation invariance applies to our perception of components within a single item, where repeated sub-sequences within a melody are always recognised as such. Thus the invariance applies to both whole melodies and portions of a melody. It applies to both speech and music. It serves the functional requirement that the same speech melody has the same meaning on different occasions, and that meanings can be assigned to portions of speech melody within an overall speech melody.

Time scaling invariance is the ability to recognise the identity of a melody or rhythm played at different tempos. It applies to whole melodies. It applies to both speech and music. It serves the functional requirement that a listener can understand speakers who speak at different speeds.

Pitch translation invariance is the preservation of musical quality when music is shifted up or down by a fixed interval. It applies to whole melodies. It applies to both speech and music. It serves the functional requirement that a listener is able to understand speech spoken by people whose voices have different pitch ranges. It is a very exact musical symmetry in the sense that the musical quality of music is almost unchanged by moderate pitch translations.

Amplitude scaling invariance is the preservation of musical quality when music is heard louder or softer. There is one aspect of music perception that is *not* amplitude scaling invariant: the intensity of the emotional and pleasurable effect of music is often increased if music is played louder, up to some level which is generally a function of how much the listener likes that particular item of music.

Octave translation invariance does not appear to serve any major functional requirement, and it is not particular evident in the perception of speech (although this may simply be because individual speech melodies do not normally contain components separated from each other by an octave). The most likely explanation for this invariance is that it provides the brain a means to economise the representation and processing of pitch value information, in that *different* pitch values separated by one or more octaves are represented by activity in the *same* neurons within certain cortical maps. This economy is likely to be particularly important for the implementation of "subtraction tables" that calculate the sizes of intervals between pairs of pitch values.

Where pitch values are reduced to pitch values modulo octaves, information about absolute pitch values is still retained, but at a lower level of precision. There is enough overlap between information in the modulo-octaves representation of pitch values and the information in the absolute representation of pitch values to allow unambiguous representation of the full absolute pitch value.

Pitch reflection invariance is a plausible symmetry, but we cannot be as certain of its existence as we can for the other five symmetries. If it does exist, it results from a symmetry of mutual interaction of neurons in certain cortical maps representing pairs of pitch values separated by certain intervals, where the interaction is a function of the interval size. Pitch reflection invariance might explain the fact that the preferred home chords A minor and C major on the diatonic white notes scale are reflections of each other about the point of reflective symmetry of the scale itself (which is the note D).

Symmetries are an important concept in the study of music for several reasons:

- If a symmetry is non-trivial to implement (as seems to be the case for pitch translation invariance and time scaling invariance), then there must be significant neural machinery involved in its implementation, and there must be a very good reason why it needs to be implemented.

- Occurrence of a given symmetry in both speech perception and music perception is strongly suggestive of a significant relationship between the two.

- The indirect manner in which some of the hypothesised musical cortical maps represent information about melody makes more sense if we realise that these cortical maps are designed to produce characterisations of input information that are invariant under the required symmetries.

- The brain may employ similar mechanisms to achieve perceptual invariance for different types of symmetry. This extends beyond music and speech: many aspects of visual perception are invariant under certain transformations, such as image translation, scaling and rotation. Understanding the perceptual invariances involved in the perception of speech and music may help us understand other types of perceptual invariance—it is likely that the brain reuses the same solution to the same problem.

16.2 Individual Cortical Maps

16.2.1 Scale Map

Assumption: The cortical map that responds to musical scales is a map that responds to recent occurrences of pitch values modulo octaves in a speech melody.

This map has the property that it develops a pattern of activity in which some areas are constantly active and other areas are inactive if, and only if, the melody is formed from notes taken from a scale which is a fixed set of pitch values modulo octaves.

Musical scales are not like any other sound stimulus that a person is normally going to be exposed to. Alternative explanations of how and why the brain perceives musical scales fall into two main groups:

- **Hard-wired:** Musical portions of the brain are presumed to exist for the purpose of listening to music. Cortical maps that respond to notes in scales exist in order to respond to music because music is constructed from scales. The problem with this theory is its obvious circularity: it presumes the existence of music played on musical scales in order to explain why we have cortical maps that respond to scales.

- **Soft-wired:** The alternative is to assume that certain cortical maps organise themselves in response to the scales that music has. This theory implicitly assumes a sufficient degree of cortical plasticity: that the brain does not expect to encounter anything like musical scales, but given that they occur it is sufficiently flexible to organise parts of itself to respond to those scales. This level of cortical plasticity can explain potentially any feature of music. The difficulty remains to explain why music takes on the attributes that it does. For example, cortical plasticity should allow the brain to perceive and process music that is *not* constructed from notes in a scale. However, such music does not occur in practice.

16.2.2 Harmonic Map

Assumption: The cortical map that responds to chords is one that is activated by pitch values that are consonantly related to pitch values already active in the map. Activity in this map tends to be reset (to zero) by a combination of a strong beat in the rhythm and a low pitch value (representing the bass) corresponding to the root note of a new chord. The purpose of this map is to detect the

occurrence of sequential (i.e. not simultaneous) pitch values that are harmonically related.

Much of the literature on harmony assumes that to understand harmony one must understand how the brain responds to simultaneous pitch values. This seems reasonable enough: the very definition of **harmony** is that it consists of simultaneous occurrence of different notes.

But if music perception is really speech perception, then there are no simultaneous pitch values to be perceived. Any cortical map that happens to be able to respond to simultaneous pitch values must have as its primary function a response to pitch values that are *not* simultaneous.

16.2.3 Home Chord Map

Assumption: There is a cortical map that characterises pitch values in a melody according to the relationships between each pitch value and the other pitch values in the same melody. The perception of home notes and home chords is a consequence of the operation of this map. The result of this perception is a characterisation of notes that is pitch translation invariant.

Because there are two strongly preferred home chords for the diatonic scale, it seems reasonable to assume that the preference for these home chords is an intrinsic function of the structure of the scale itself. Since home chords are pitch translation invariant, the processes that determine the preferred home chord for a scale must be pitch translation invariant, and therefore must be a function of the intervals between pitch values in the scale. The most likely candidates are (1) proximity of pitch values (i.e. small intervals) and (2) consonance of intervals.

Assumption: The preference for home chord in a scale is largely a function of consonance relationships between notes on the scale.

If we define the **neighbourhood** of a note to be the pattern of notes close to that note, it is observed that notes with similar neighbourhoods do not have similar levels of "home" quality. For example, comparing C and F in the white notes scale, the neighbourhood of C consisting of G, A, B, C, D and E is an exact transposition of the neighbourhood of F consisting of C, D, E, F, G and A. But the "home" quality of the two notes is quite different, as C is the root note of the possible home chord C major, whereas F does not belong to any possible home chord on the scale.

However, if the notes in the white notes scale are unfolded into the **Harmonic Heptagon**, then there is an obvious relationship between the sets of notes that are most strongly connected to other notes by consonant relationships, and the sets of notes that form the preferred home chords. In particular, the notes B and F have one less connection to other notes, and D is connected to both of these. Assuming a pattern of mutual reinforcement via connections to other notes in the heptagon, this leaves A, C, E and

G as being the most strongly reinforced, and these happen to be the notes contained in the two preferred home chords.

To explain why the home chord has to be *either* A minor or C major, we have to assume that there is some reason why the home chord cannot be A minor 7th, i.e. ACEG. The simplest explanation is that the home chord cortical map disallows activation of notes not consonantly related, i.e. A and G, so that one or the other of those two has to be left out.

16.2.4 Regular Beat Map

Assumption: There is a cortical map that contains a map of neurons activated by the occurrence of regular beats of fixed tempos. These neurons tolerate omissions of individual beats (implying they maintain their own "internal" beat), and they do not respond to beats occurring out of phase. The purpose of this map is to respond to the *irregular* rhythms of speech.

The regular beat map responds to the basic musical feature of regular hierarchical time structure.

16.2.5 Note Duration Map

Assumption: There is a cortical map that contains a map of neurons activated by the lengths of the durations between beats occurring within a rhythm.

16.2.6 Melodic Contour Map

Assumption: There is a cortical map that contains a map of neurons activated by the rising and falling of pitch values in a melody.

16.3 Repetition

Assumption: Repetition is an explicit aspect of music, and this implies that there are specific mechanisms in the brain that play a role in the perception of repetition. There is a distinction between the perception of free and non-free repetition, where a repetition count is maintained only in the case of non-free repetition.

16.4 Assumptions of the Final Theory

16.4.1 General Principle of Music

Assumption: There exists a general principle of musicality.

There are many different aspects of music, i.e. melody, harmony, bass, home notes, home chords, rhythm, tempo and repetition. Different types of music emphasise different aspects to different extents. It is possible to perceive musicality even when some aspects are completely missing. The best example of this is the musicality of purely percussive music, which has temporal aspects but no pitch-related aspects (all rhythm, no melody).

This aspectual nature of music suggests that there may be some general principle of musicality which is perceived separately in each cortical map that responds to a different aspect of music, such that the musicality perceived in different cortical maps is then combined to calculate an overall musicality.

16.4.2 Echoing

Assumption: The listener can detect certain patterns of neural activity in cortical maps in the speaker's brain involved in the generation of speech, echoed in patterns of neural activity in their own corresponding cortical maps when they perceive speech.

The assumption is that there is some correspondence between neural activity when you generate speech and neural activity when you listen to the same speech spoken by someone else.

16.4.3 General Principle and Conscious Arousal

Assumption: The general principle of musicality reflects some aspect of the internal mental state of the speaker.

The echoing assumption allows us to connect the perception of patterns of activity in the listener's own brain to the perception of patterns of activity in the speaker's brain.

Conscious arousal is assumed to be a general aspect of mental state which causes an alteration in the mode of neural activity in large areas of the brain, and which is modulated by means of **non-specific neurotransmission**, where certain neurotransmitters are released non-specifically into the brain medium such that their action is distributed over neurons in a large area. It is assumed that this modulation causes an alteration in certain average characteristics of neural activity, and this characteristic is echoed in the brain of the listener, and the echoed characteristic can be measured by those neurons detecting musicality.

16.4.4 Constant Activity Patterns

Assumption: The specific patterns of neural activity corresponding to perceived musicality are constant zones of activity and constant zones of inactivity, with a maximal border perimeter between the two.

This assumption is consistent with the distinct zones of activity and inactivity that occur in both the scale cortical map and the regular beat cortical map, in response to music, but which do not occur in response to the smooth melodic contours and irregular rhythms of normal speech. These **constant activity patterns** can be detected by edge-detecting neurons which respond to the occurrence of persistent edges between zones of active neurons and zones of inactive neurons in a cortical map.