Psychology and Music

Diana Deutsch
University of California, San Diego

INTRODUCTION

The relationship between psychology and music is characteristic of that between a new science and an established discipline. Western music theory has a very old tradition, dating at least from the time of Pythagoras; and the philosophical underpinnings of this tradition that were established in ancient times still exist today. Most characteristic of this tradition is its rationalism. In contrast with the scientific disciplines, the development of music theory over the last few hundred years has not been characterized by a growth in the empirical method. Rather, while composers have constantly experimented with new means of expression, music theorists have on the whole been system builders who sought to justify existing compositional practice or to prescribe new practice on numerological grounds. Further, when an external principle has been invoked as an explanatory device, most commonly such a principle was taken from physics. The concept of music as essentially the product of our processing mechanisms and therefore related to psychology has only rarely been entertained.

There are several reasons why this rationalistic stance was adopted, most of which no longer apply. One reason was a paucity of knowledge concerning the nature of sound. It is understandable that the inability to characterize a physical stimulus should have inhibited the development of theories concerning how this stimulus is processed. A related reason was poor stimulus control, which made experimentation difficult. A third reason was the lack of appropriate mathematical techniques with which to study probabilistic phenomena. However, another reason, which is still with us today, lies in the peculiar nature of music itself. There are no external criteria for distinguishing between music and nonmusic, or between good music and bad music. Further, it is clear that how we perceive music depends at least to some extent on prior experience. Thus the relevance of psychological experimentation to music theory requires careful definition.

In this chapter I first review major developments in music theory from an historical point of view. Following this I explore various issues that are currently being studied both by music theorists and by psychologists. Finally, I discuss the role of psychology in music theory.

HISTORICAL PERSPECTIVE

Speculations concerning music may be traced back to very ancient times (Hunt, 1978), but the foundations of Western music theory are generally held to have been laid by Pythagoras (ca. 570-497 B.C.). Pythagoras was concerned mostly with the study of musi-

cal intervals. He is credited with identifying the musical consonances of the octave, fifth, and fourth with the numerical ratios 1:2, 2:3, and 3:4. He is also credited with establishing by experiment that the pitch of a vibrating string varies inversely with its length. However, Pythagoras and his followers ultimately lost faith in the empirical method and instead attempted to explain all musical phenomena purely in terms of numerical relationships. As Anaxagoras (ca. 499-428 B.C.) declared: "Through the weakness of the sense-perceptions we cannot judge truth [Freeman, 1948, p. 86]." And later Boethius, the leading music theorist of the Middle Ages and a strong follower of Pythagoras, wrote in *De Institutione Musica:*

> For what need is there of speaking further concerning the error of the senses when this same faculty of sensing is neither equal in all men, nor at all times equal within the same man? Therefore anyone vainly puts his trust in a changing judgement since he aspires to seek the truth [Boethius, 1967, p. 58].

The view that music ought to be investigated solely by contemplation of numerical relationships has characterized most music theory since Pythagorean times. On this view, the world of mathematics is held to provide an ideal which the world of sense-perception can only imitate. Experimental procedures are therefore held to be irrelevant: if the results of experiments are in accordance with theory, then they are redundant; if the results conflict with theory, then they must have been ill-conceived in the first place. Also stemming from the mathematical approach of the Pythagoreans have been the numerous attempts to build entire musical systems by mathematical deduction from a minimal number of established musical facts. Essentially this approach derives from a false analogy with geometry (Russell, 1945). Euclidean geometry begins with a few axioms which are held to be self-evident, and from these axioms arrives by deduction at theorems that are not in themselves self-evident. However, it is a logical error to assume that we can proceed *by deduction* from one musical fact to another musical fact. Properly, musical facts can only be used as a basis for the formulation of *hypotheses* about further musical facts, which require empirical verification.

Another strong influence on music theory which stemmed from the Pythagoreans was the belief that the ultimate explanation of musical phenomena lies in physics. Until the Copernican revolution, this belief took the form of assuming that music serves as a reflection of sounds produced by the heavenly bodies. As described by Aristotle in *De Caelo*, it was thought:

> that the motion of bodies of that [astronomical] size must produce a noise, since on our earth the motion of bodies far inferior in size and in speed of movement has that effect. Also, when the sun and the moon, they say, and all the stars, so great in number and size, are moving with so rapid a motion, how should they not produce a sound immensely great? Starting form this argument, and from the observation that their speeds, as measured by their distances, are in the same ratio as musical concordances, they assert that the sound given forth by the circular movement of the stars is a harmony [Aristotle, 1930, p.290].

Figure 1 shows that Pythagorean view of the universe, in which the relative distances of the heavenly bodies to each other are displayed, together with the musical intervals formed thereby. It can be seen that the distance between the Earth and the Moon formed

a whole tone, from the Moon to Mercury a semitone, from Mercury to Venus a semitone, from Venus to the Sun a tone and a half, from the Sun to Mars a whole tone, from Mars to Jupiter a semitone, from Jupiter to Saturn a semitone, and finally from Saturn to the Supreme Heaven, a semitone. Notice further that the entire distance between Earth and the Supreme Heaven formed an Octave.

The theory of the Harmony of the Spheres was an attractive one, since it provided answers to several fundamental questions about music. One question was why music exists in the first place; and the answer provided was that it serves as a reflection of the Divine Harmony. A second question was why certain musical intervals (the consonances) strike us as pleasing while others do not; and the answer here was that the consonances are those intervals that are present in this Divine Harmony. The theory even had a normative value, since it provided boundary conditions for separating music from non-music.

The main problem with the theory that puzzled the ancient Greeks (as well as those who followed) was why, if the heavenly bodies do indeed produce this harmony, we cannot hear it. One answer, suggested by Censorinus, was that the loudness of the sound is so great as to cause deafness[1] (Hawkins, 1853/1963). An alternative view, described by Aristotle (who did not in fact endorse it), was that since this sound is with us since birth, and since sound is perceived only in contrast to silence, we are not aware of its presence. However, neither of these views was considered satisfactory.

At all events, the theory of the Harmony of the Spheres provided a strong link among the studies of music, astronomy, and mathematics, with the result that the scientific part of the program of higher education developed into the *Quadrivium* of the "related studies" of astronomy, geometry, arithmetic, and music. The Quadrivium persisted through to the end of the sixteenth century and was responsible for much interaction between the disciplines.
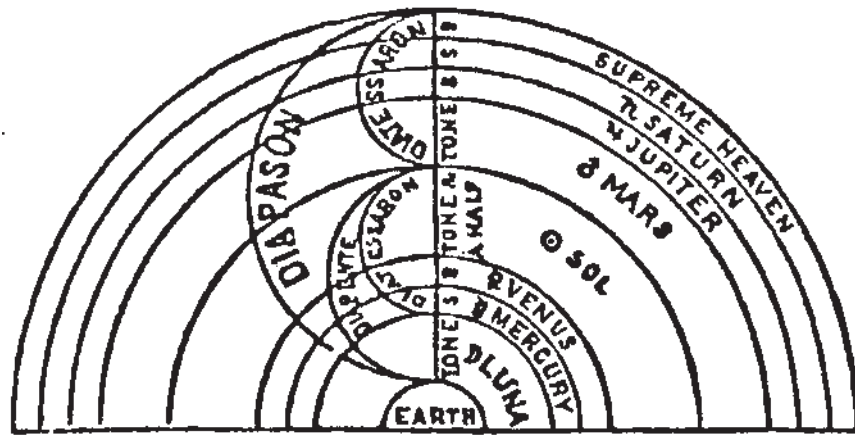

FIG. 1. Pythagorean view of the universe in musical intervals. (From Hawkins, 1853/1963.)

[1] This view inspired Butler's lines in *Hudibras* (Part II):

> Her voice, the music of the spheres,
> So loud it deafens mortal ears,
> As wise philosophers have thought,
> An that's the cause we hear it not.
> [Butler, 1973, p. 122].

In general, the later Greek theorists adhered to the numerological approach of the Pythagoreans. There was, however, a notable exception. Aristoxenus (ca. 320 B.C.), originally a pupil of the Pythagoreans and later of Aristotle, saw clearly that music cannot be understood by contemplation of mathematical relationships alone. He argued that the study of music should be considered an empirical science and that musical phenomena were basically perceptual and cognitive in nature. For example, in the *Harmonic Elements* he wrote:

> The order that distinguishes the melodious from the unmelodious resembles that which we find in the collocation of letters in language. For it is not every collocation but only certain collocations of any given letters that will produce a syllable.

And later:

> It is plain that the apprehension of a melody consists in noting with both the ear and intellect every distinction as it arises in the successive sounds-successive, for melody, like all branches of music, consists in a successive production. For the apprehension of music depends on these two faculties, sense-perception and memory; for we must perceive the sound that is present and remember that which is past. In no other way can we follow the phenomenon of music [Aristoxenus, 1902, pp. 192-194].

But Aristoxenus was not understood by his contemporaries, nor by the music theorists of the Middle Ages and early Renaissance, who continued to adhere to the numerological approach. Most of his works were lost to posterity, though fortunately two books of his *Harmonic Elements* and fragments of his *Elements of Rhythmics* were preserved.

In violation of prevailing theoretical constraints, medieval polyphony employed intervals other than the pure consonances of the octave, fifth, and fourth allowed by the Pythagoreans. It therefore fell to the theorists of the fifteenth and sixteenth centuries to justify existing practice in the context of the Pythagorean doctrine. This was achieved by Zarlino (1517-1590) who argued that the number six had various metaphysical properties. For example, it is the first perfect number ($1 + 2 + 3 = 1 \times 2 \times 3 = 6$). Zarlino proposed that the realm of the consonances be extended to combinations produced by ratios formed by the first six numbers. This justified the use of the major third (5:4), minor third (6:5), and major sixth (5:3). (The minor sixth was also admitted somehow, although its ratio is 8:3.) In his heavily numerological and theological treatise, *Istituzioni Armoniche* (1558/1950), Zarlino developed rules of composition based on the concept of the first six numbers as a divinely ordained sanctuary containing the consonances (the *scenario*) outside of which the composer can wander only under severe restrictions. Thus theoretical approval was given to existing musical practice on numerological grounds, and a new set of boundary conditions for music was established (Palisca, 1961).

The scientific revolution of the sixteenth and seventeenth centuries had a profound effect on music theory. First, advances in astronomy forced theorists to abandon the view that the universe was a harmony, and with it the view that musical consonances reflect this harmony. Second, advances in understanding the properties of vibrating strings led to a re-evaluation of the role of number in musical explanation: Numerical ratios were now considered meaningful in that they applied to the properties of sounding bodies. Discovery of the overtone series, of the relationship between pitch and frequency, and of

the physical correlates of consonance and dissonance inclined some thinkers to adopt a more empirical approach to musical issues in general (Palisca, 1961).

Notable among the musical empiricists of the sixteenth century were Giovanni Battista Benedetti (1530-1590) and Vincenzo Galilei (1520-1591).[2] Benedetti was perhaps the first to relate the sensations of pitch and consonance to rates of vibration. Galilei demonstrated by experiment that the association of the consonant intervals with simple numerical ratios held only when their terms represented pipe or string lengths and also when other factors were held constant. For example, these relationships did not hold for relative weights of hammers, nor for volumes enclosed in bells. He also argued that disputes concerning tuning systems were futile, since the ear cannot detect the small pitch differences under debate. He proposed a new theory of counterpoint based on existing musical practice, rather than on appeal to extra-musical phenomena, and he argued strongly for the empirical method in studying music. However, thinkers such as Galilei were very much in the minority, and the prevailing theoretical stance continued to be heavily rationalistic.

In parallel with scientific advances concerning the physical properties of sound[3], composers of the late sixteenth and the seventeenth centuries were particularly active in experimenting with new techniques. There thus arose a need for a new theoretical synthesis to justify prevailing musical practice and to link this with newly obtained scientific knowledge. This was achieved by the composer and music theorist Jean-Philippe Rameau (1683-1764). Rameau's systematization forms the basis of traditional harmonic theory as we know it today. By analyzing the compositions of his predecessors and contemporaries and by joining to these analyses the results of his own musical investigations, Rameau arrived at important fundamental laws and concepts such as the invertibility of chords, the generation of a chord by its root, the root progression chords, and so on.

In one sense, Rameau's synthesis can be regarded as a great psychological achievement, in which he used as his body of data the music of common practice to formulate a viable theory of the abstract structure of music. However, Rameau did not regard music as essentially the product of our perceptual and cognitive mechanisms; rather, true to tradition, he felt the need to justify his system in terms of a single physical principle. He found this in the recently discovered phenomenon of the overtone series, and so he invoked it as the "self-evident principle" from which he attempted to derive an entire musical system by mathematical deduction. As he wrote:

> Music is a science which ought to have certain rules; these rules ought to be derived from a self-evident principle; and this principle can scarcely be known to us without the help of mathematics [Rameau, 1722/1950, p.566].

Although his attempts to manipulate the numerical ratios failed and involved him in a mass of inconsistencies and contradictions, Rameau's approach laid the groundwork for

---

[2] Galilei was the father of Galileo

[3] Many noted scientists of the seventeenth century addressed themselves to issues concerning sound and music, notably in the areas of pitch and interval relationships. These included Galileo, Mersenne, Descartes, Kepler, and Huygens.

a new musical numerology in which the overtone series replaced the Harmony of the Spheres as the ultimate explanatory device (Palisca, 1961).

Perhaps the greatest music theorist of the nineteenth century was Hermann von Helmholtz (1831-1894), whose book *On the Sensations of Tone* (1885/1954) makes important reading even today. Helmholtz saw clearly that musical phenomena require explanation in terms of the processing mechanisms of the listener. He carried out important experimental work on issues such as the perception of pitch, combination tones, beats, and consonance and dissonance. He also speculated concerning the nature of high-level cognitive mechanisms underlying music perception, though he lacked the technical resources to investigate these mechanisms experimentally.

Technological advances of the end of the last century and the beginning of this one enabled scientists for the first time to investigate auditory phenomena under strictly controlled conditions (see Marks, in Volume III). The science of psychoacoustics was thus established. However, the sound stimuli that could be precisely generated were very limited in scope. It became possible, for example, to perform careful measurements on auditory threshold phenomena and to devise psychophysical scales of pitch and loudness. However, it was still prohibitively difficult to construct sequences of tones under controlled conditions or to generate tones with specified time-varying spectra. Thus, the issues to which psychoacousticians addressed themselves were not of much concern to musicians, who found the perceptual properties of simple auditory stimuli in isolation of little theoretical interest.

Matters were made worse by certain conclusions from psychoacoustics which musicians felt were at variance with their experience and intuitions. One notable example is the mel scale for pitch (Stevens & Volkmann, 1940). As shown on Figure 2, this scale designates as equal, intervals which are unequal on the musical scale; and conversely, equal musical intervals are designated as unequal on the mel scale. Thus, it seemed to many musicians that, however carefully controlled the psychoacoustical experiments were, they were leading to incorrect conclusions. Rather than criticizing these conclusions on home ground, musicians regarded them as evidence that scientific methodology was inappropriate for the study of music.

At the same time as the science of psychoacoustics was developing with its focus on narrow stimulus parameters, music theorists were finding themselves faced with a vast increase in the complexity of the music that they were attempting to explain. The development of chromaticism in the music of the nineteenth and early twentieth centuries, for example in the music of Wagner, Debussy, Moussorgsky, and Mahler, forced a fundamental change in the concept of harmony. First the concept of tonality developed into the concept of extended tonality to accommodate these new complexities. However even this latter concept had to be abandoned, since it became dubious whether the notion of a tonic served as a useful explanatory concept for the new compositions. Music theorists therefore began to search for an entirely new theoretical framework within which they could compose.

The framework which became the most influential was the twelve-tone system, originally developed by Schoenberg. This system, which is described below, has inspired much theoretical work on equivalence relations between sets of pitches. However,
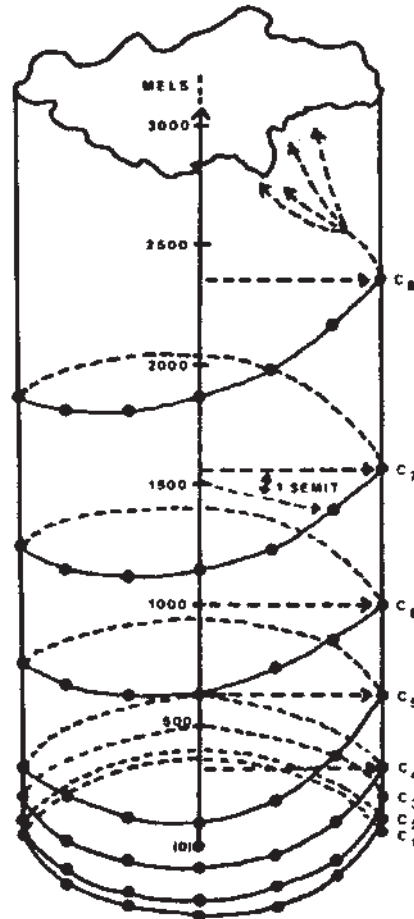
FIG. 2. Pitch as scaled in mels and in octaves. (From Ward & Burns, 1982.)

twelve-tone theorists did not deem it appropriate to determine experimentally whether the equivalence relations of their system were perceptually relevant. Rather, in line with Pythagorean tradition, they considered the intrinsic plausibility of the basic axioms of the system, together with its internal consistency, as sufficient justification for its use in compositional practice.

Just as the technological advances of the first part of this century tended to create a rift between scientists and musicians, so have recent technological advances over the last decade created an era of collaboration between the disciplines. With the aid of computer technology, psychologists are now able to generate complex auditory stimuli with precision and so to examine musical issues in a controlled experimental setting. At the same time, composers have been increasingly interested in the computer as a compositional tool. However, in order to make effective use of this new technology, they need to obtain answers to questions in perceptual and cognitive psychology. As a result of these developing interests from both disciplines, there is not only a rapid expansion of empirical work on music perception and cognition, but, perhaps more importantly, increasing collaboration between psychologists and musicians. We can confidently predict that over the next decade psychology will have a firmly established place in the music theory.

SOME CURRENT ISSUES

I now turn to consider various issues concerning music perception and cognition that are currently being studied both by music theorists and by psychologists. These are likely to be the focus of future work. This review is not intended to be exhaustive, but rather illustrative of the ways in which findings from psychology can usefully be applied to music.

Music and Composed Sounds

In the music of the seventeenth, eighteenth, and early nineteenth centuries, the timbre or sound quality of an instrument was generally treated as a carrier of melodic motion, rather than as a primary compositional attribute in itself. However, the decline of tonality opened the way for new compositional uses of timbre. Composers began experimenting with complex sound structures that resulted from several instruments playing simultaneously, such that the individual instruments lost their identifiability and fused to produce a single sound impression. Debussy in particular made extensive use of chords that approached timbres. Early in this century composers such as Schoenberg, Webern, Stravinsky, and particularly Varese frequently employed such highly individualized sound structures, termed by Varese "sound masses." Such experimentation led composers to explore the characteristics of sound that were conducive to perceptual fusion (Erickson, 1975, 1982).

Developing interest in musical timbre also led composers to experiment with sound sequences involving rapid timbral changes. Such sequences, know as *Klangfarbenmelodien*, or melodies composed of timbres, were used early in this century by composers such as Schoenberg and Webern, and later by composers such as Boulez. This led to speculation concerning the rules governing orderly transitions between timbres. As Schoenberg (1911) wrote:

> If it is possible to make compositional structures from sounds which differ according to pitch, structures which we call melodies, sequences producing an effect similar to thought, then it must be possible to create such sequences from the timbres of that other dimension from what we normally and simply call timbre. Such sequences would work with an inherent logic, equivalent to the kind of logic which is effective in melodies based on pitch. All this seems a fantasy of the future, which it probably is. But I am firmly convinced that it can be realized [470-471].

In essence, Schoenberg was proposing that timbres are psychologically represented in an orderly fashion and that the structure of this representation can be exploited compositionally.

Interest in understanding the psychological representation of timbre was accelerated by the development of electronic and computer music (Matthews, 1969). With the aid of new technology, composers became able for the first time to generate any sounds they wished, free from constraints imposed by the physics of natural instruments or by the capabilities of the human performer. But this very freedom presented fundamental problems in perceptual psychology which required solution. As the music theorist and composer Robert Erickson (1975) wrote:

> A composer who wishes to carve out certain sounds from this infinity of possibilities must decide: which ones? He may attempt to create "an instrument," meaning some sort of unified selection of

sounds from the infinity of possibilities….Or he may go at things more abstractly, thinking in terms of contrast, similarly, sound classes….It may be true that we are on the edge of being able to produce any sound we can imagine, just as it is true that we can produce any pitch we can imagine. The infinity of sounds in the universe may be objectively real to physics and measuring instruments; if it is unrealizable in music then the difficulty must be related to human limitations and to the limitations imposed by musical discourse [p. 9].

Three related questions concerning timbre perception are here examined. First, what are the acoustical parameters underlying perception of instrumental timbre? Second, what parameters give rise to the perception of unitary sound images, and what give rise to the perception of multiple simultaneous sound images? Third, how do timbres behave when juxtaposed in time? It is clear that these questions all have implications not only for music, but also for auditory perception in general.

The identification of timbre. It is remarkable that the sound of a musical instrument can be identified under a wide range of conditions, regardless of its pitch, its loudness, and so on. The sound spectrograms produced by the same instrument under different conditions vary considerably. What are the features underlying such perceptual constancy?

Classically, the issue of timbre perception has been concerned with tones in the steady state. According to Helmholtz (1885/1954), differences in the timbre of complex tones depend on the strengths of their various harmonics. He claimed that simple tones sound pleasant, but dull at low frequencies; complex tones whose harmonics are moderately strong sound richer but still pleasant; tones with strong upper harmonics sound rough and sharp; and complex tones consisting only of odd harmonics sound hollow. More recently, Plomp and his collaborators have argued that the critical band[4] plays an important role in timbre perception (Plomp, 1964, 1970; Plomp & Mimpen, 1968). Evidence was obtained that harmonics falling within the same critical band fuse in their effect. Other experiments have been addressed to the question of whether perceived timbre is based on the relationships formed by the fundamental frequency and the frequency region of a formant, or on the absolute level of the formant. In general the results favor a modified fixed-formant model of timbre perception (Plomp & Steeneken, 1971; Slawson, 1968).

Recently, the investigation of timbre has concerned itself with tones produced by natural instruments. Such tones are held to consist of three temporal segments: the attack, the steady state, and the decay. The attack segment has been found to be of particular importance to timbre identification (Berger, 1964; Grey, 1975; Saldanha & Corso, 1964; Wedin & Goude, 1972; Wessel, 1973, 1978); the steady state segment contributes more to timbre identification if it varies in time; and the decay segment appears of little consequence (Saldanha & Corso, 1964).

An important technique in the study of timbre perception was pioneered by Risset and Matthews (1969). Samples of natural instrument tones are digitized and analyzed by computer, a set of physical parameters is extracted from this analysis, and tones are then synthesized by computer in accordance with these parameters. This technique enables the experimenter to vary systematically any parameters, and so to examine the

4 The critical band is that frequency band within which the loudness of a band of sound of constant sound pressure level is independent of bandwidth.

perceptual effects of these variations. It has been shown using this technique, for instance, that when tones are resynthesized with a line-segment approximation to the time-varying amplitude and frequency function for the partials, there is very little loss of characteristic perceptual quality, though considerable information reduction may thus be produced (Grey & Moorer, 1977).

Also using this technique, geometric models of subjective timbral space have been generated. Instrument sounds that are judged as similar are positioned close together in this space; sounds that are judged as dissimilar are positioned far apart. Such models have been provided by Wessel (1973, 1978) and by Grey (1975) for string and wind instrument tones that were equated for pitch, loudness, and duration. At least two dimensions have been unveiled: The first appears to relate to the spectral and distribution of sound energy, and the second to temporal features such as details of the attack.

With such representations, it has proved possible to draw trajectories through a given timbral space and so to create interpolated sounds that are consistent with the geometry of the space. For example, Grey (1975) created a series of tones which traversed his multidimensional space in small steps, so that the listener first perceived one instrument (such as a clarinet) and at some point in the series realized that he was now hearing a different instrument (such as a cello). Yet the perceptual transition between instruments appeared completely smooth. Thus Schoenberg's vision of composing with timbres that are arranged along an orderly continuum appears realizable. However, before these models can be used flexibly they will require considerable elaboration to accommodate the invariance of timbre under pitch and loudness changes, as well as effects of context. (See also Risset & Wessel, 1982.)

Spectral fusion and separation. A fundamental task for auditory theory is to define the relationships between components of an ongoing acoustic spectrum that result in the perception of a unitary sound image, and those that result in the perception of several simultaneous but distinct sound images. These processes of fusion and separation are of basic importance, since without them there would be no intelligible listening at all. Presumably, we have evolved mechanisms that lead us to fuse together elements of the spectrum that are likely to be emanating from the same source, and to separate out those that are likely to be emanating from different sources. This view of perception as a process of "unconscious inference" was originally proposed by Helmholtz (see Helmholtz, 1909-1911/1925) and has recently been invoked to explain various findings in perceptual psychology, both in vision (e.g., Gregory, 1970; Hochberg, 1974; Sutherland, 1973) and in hearing (e.g., Bregman, 1978; Deutsch, 1975a, 1979; Warren, 1974).

With specific regard to music, Helmholtz (1885/1954) posed the question of how, given the rapidly changing, complex spectrum resulting from several instruments playing simultaneously, we are able to reconstruct our musical environment so that some components of the spectrum give rise to a unitary sound image, and others give rise to several distinct but simultaneous sound images. Thus, he wrote:

> Now there are many circumstances which assist us first in separating the musical tones arising from different sources, and secondly, in keeping together the partial tones of each separate source. Thus when one musical tone is heard for some time before being joined by the second, and then

the second continues after the first has ceased, the separation in sound is facilitated by the succession of time. We have already heard the first musical tone by itself, and hence know immediately what we have to deduct from the compound effect for the effect of this first tone. Even when several parts proceed in the same rhythm in polyphonic music, the mode in which the tones of different instruments and voices commence, the nature of their increase in force, the certainty with which they are held and the manner in which they die off, are generally slightly different for each…but besides all this, in good part music, especial care is taken to facilitate the separation of the parts by the ear. In polyphonic music proper, where each part has its own distinct melody, a principal means of clearly separating the progression of each part has always consisted in making them proceed in different rhythms and on different divisions of the bars.

And later:

All these helps fail in the resolution of musical tones into their constituent partials. When a compound tone commences to sound, all its partial tones commence with the same comparative strength; when it swells, all of them generally swell uniformly; when it ceases, all cease simultaneously. Hence no opportunity is generally given for hearing them separately and independently [Helmholtz, 1885/1954, pp. 59-60].

One factor proposed by Helmholtz as promoting fusion was onset synchronicity of spectral components. This has recently been shown to be important in several studies. Rasch (1978) investigated the threshold for perception of a high tone when this was accompanied by a low tone. He found that when the onset of the low tone was delayed relative to the high tone there was a substantial lowering of threshold. In addition, the percept when the tones were asynchronous was very different from the percept when the tones were synchronous; in the former case, two distinct tones were clearly perceived, but in the latter case, they fused to produce a single percept. Bregman and Pinker (1978) employed a paradigm in which a simultaneous two-tone complex was presented in alternation with a third tone. With increasing asynchrony between the simultaneous tones there was an increase likelihood that one of these would form a melodic stream with the third tone. Both sets of authors interpret their findings along the lines advanced by Helmholtz. A related study on the effects of asynchrony was performed by Deutsch (1979) using spatially separated tones (see p. 15).

A second factor proposed by Helmholtz to promote fusion is coordinated modulation in the steady state. McNabb and Chowning have shown informally that with a harmonic tone complex whose spectrum corresponds to a vowel the impression of a voice is strongly enhanced when a small amount of coordinated frequency modulation, which can be either periodic (vibrato) or random (shimmer), is superimposed on all components simultaneously. McAdams and Wessel have informally investigated the effect of imposing two different modulation functions on the odd or even partials of a complex tone and report that this produced the impression of two simultaneous sounds (see McAdams, 1981).

A third factor that has been hypothesized to promote fusion is harmonicity of the components of a complex spectrum. Stringed and blown instruments, which tend to produce strongly fused images, have partials that are harmonic or nearly harmonic. However bells and gongs, which produce diffuse images, have partials that are nonharmonic (Matthews & Pierce, 1980). DeBoer (1976) has shown that harmonic complexes

tend to produce unitary and unequivocal pitch sensations, whereas various kinds of non-harmonic complexes produce multiple pitch sensations. Again, this is expected on the assumption that our auditory mechanisms have evolved so as to make the most proba-ble interpretations in terms of sound sources, since most forced vibration systems such as the voice have partials whose frequencies are harmonic or close to harmonic.

Perception of sequences of timbres. As noted above, twentieth-century composers have become interested in the production of sound sequences involving rapid changes of tim-bre. This raises the question of how sequences of contrasting timbres are perceived. An effect of central interest here was first reported by Warren, Obusek, Farmer, and Warren (1969) in a paper entitled (rather ironically): "Auditory sequence: Confusions of patterns other than speech or music." These authors constructed repeating sequences of four unrelated sounds: a high tone (1000 Hz), a hiss (2000 Hz octave band noise), a low tone (796 Hz), and a buzz (4000 Hz square wave). Each sound lasted for 200 msec, and the different sounds followed each other without pause. Listeners were found to be quite unable to name the orders of such repeating sounds. The duration of each sound had to be increased to over 500 msec for correct ordering to be achieved.

The "Warren effect" probably has two bases. The first is that listeners tend to organ-ize sounds into separate streams on the basis of sound type; and auditory streaming pro-duces difficulty in forming temporal relationships across streams (see p. 173). Indeed, the threshold for ordering two acoustic events is higher when these events are dissimilar than when they are similar (Hirsh, 1959; Hirsh & Sherrick, 1961). Second, Warren (1974) has hypothesized that unfamiliarity with such a sound sequence contributes to difficul-ty in ordering. At all events, this type of study shows that with rapid contrasting sounds the listener may be unable to obtain the impression of a coherent sequence and may instead perceive multiple sequences in parallel.

Another effect of context was studied by Bregman and Pinker (1978). In conditions where a two-tone complex alternates with a third tone, if one of the tones in the complex is similar in frequency to this third tone, this component may detach itself perceptually so as to form a melodic stream with the third tone. When this happens there is an alter-nation in perceived timbre for the two-tone complex. Thus, the timbre of any given sound is likely to vary depending on the sequential context in which this sound is embedded.

In summary, the study of timbre perception is a particularly good example of fruitful collaboration between psychologists and musicians. Most of the questions so far raised in this area have been by musicians who were concerned with solving compositional problems; however, these questions are fundamental to the understanding of sound per-ception in general. Progress toward answering these questions probably could not have been achieved without the use of experimental techniques developed by psychologists.

Music and the Performing Space

Composers have long been concerned with spatial aspects of music; however interest in this area has developed particularly since Berlioz (1806-1869) who argued that the dispo-

sition of instruments in space should be considered an essential part of a composition. In his *Treatise on Instrumentation*, Berlioz wrote:

> I want to mention the importance of the different points of origin of the tonal masses. Certain groups of an orchestra are selected by the composer to question and answer each other; but this design becomes clear and effective only if the groups which are to carry on the dialogue are placed at a sufficient distance from each other. The composer must therefore indicate in his score their exact disposition. For instance, the drums, bass drums, cymbals and kettledrums may remain together if they are employed, as usual, to strike certain rhythms simultaneously. But if they execute an interlocutory rhythm, one fragment of which is given to the bass drums and cymbals, the other to kettledrums and drums, the effect would be greatly improved and intensified by placing the two groups of percussion instruments at the opposite ends of the orchestra, i.e., at a considerable distance from each other [Berlioz, 1948, p. 407].

Later composers such as Ives, Brant, and Stockhausen paid particular attention to the positioning of instruments and instrument groups and carried out informal experiments to investigate the effects of different spatial arrangements on the way music is perceived (see, e.g., Brant, 1966).

In a controlled experimental setting, spatial relationships have been shown to interact with other musical attributes in systematic ways. Earphone listening provides a particularly well-defined situation for examining the effects of spatial separation; and results obtained under these conditions can later be tested for generality in free sound-field environments (Deutsch, 1982a).

Deutsch (1975a, 1975b) examined the perceptual effects of presenting two simultaneous sequences of tones, one to each ear. The following question was raised. Does the listener, under these conditions of extreme spatial separation, perceive the sequence emanating from one side of space or the other; or, does the listener instead form perceptual configurations on a different basis?

The stimulus pattern employed to examine this issue is shown in Figure 3a. It consisted of a major scale, presented simultaneously in both ascending and descending form, such that when a tone from the ascending scale was in one ear, a tone from the descending scale was in the other ear, and successive tones in each scale alternated from ear to ear. No listener perceived the sequence of tones presented to one side of space or to the other. Instead, most listeners obtained the percept shown in Figure 3b. This consisted of two melodic lines, one formed by the higher tones and the other by the lower tones. Further, the higher tones all appeared to emanate from one earphone, and the lower tones from the other. A minority of listeners perceived instead a single melodic line that corresponded to the higher tones, and they perceived little or nothing of the lower tones. Thus for all listeners, the formation of perceptual configurations on the basis of pitch proximity was so strong as to override completely the effects of spatial separation and often to produce striking localization illusions. The tones were perceptually reorganized in space to be consistent with pitch proximity.

Further findings concerned localization patterns for the higher and lower tones, and their handedness correlates. Righthanders showed a pronounced tendency to hear the higher tones as on the right and the lower tones as on the left, regardless of their true locations. However, lefthanders did not show this tendency. Since the left hemisphere
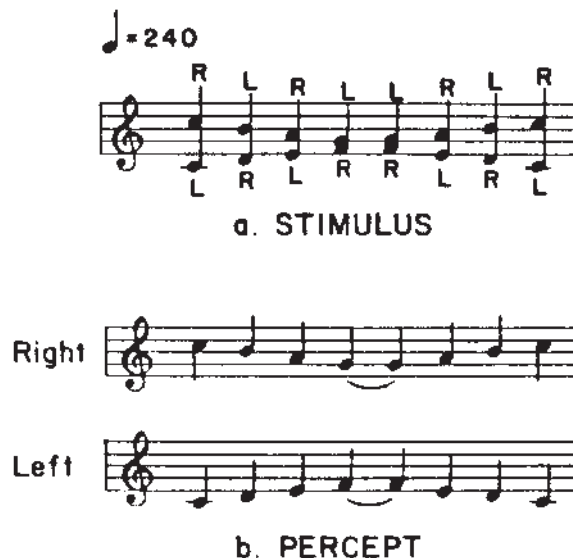
FIG. 3. (a) Configuration giving rise to the scale illusion. (b) Illusion most commonly produced. (From Deutsch, 1975a.)

is dominant in most righthanders, this pattern of results indicates that we tend to hear the higher tones as coming from the side of space that is contralateral to the dominant hemisphere, and the lower tones as from the other side (Deutsch, 1975a, 1975b).

This study was followed up by the music theorist Butler (1979a) who was concerned with determining the generality of these findings in natural musical situations. He presented the scale configuration through loudspeakers in a free sound-field environment and asked music students to notate separately the sequence that they heard coming from the speaker on the right and the sequence that they heard coming from the speaker on the left. In some conditions piano tones were used as stimuli. Despite these differences, essentially the same pattern of results emerged: Virtually all listeners heard the higher tones as emanating from one speaker and the lower tones as from the other. The effects were also explored of introducing differences in loudness and timbre between the stimuli coming from the two speakers. This resulted in a change in tone quality, however the new sound was heard as though coming simultaneously from both speakers. Thus, not only were the spatial locations of the tones perceptually rearranged to accommodate pitch proximity, but their timbres and loudnesses were rearranged also. Butler also devised different contrapuntal patterns which were played to listeners through earphones or spatially separated loudspeakers. Essentially the same results were obtained: The patterns were perceptually reorganized so that a higher melodic line appeared to be emanating from one earphone or speaker, and a lower melodic line from the other.

Such effects are found in performed music. For example, the last movement of Tschaikowsky's Sixth Symphony (the "Pathetique") begins with a passage in which the theme and accompaniment are distributed between two violin parts. However, the theme is heard as coming from one set of violins and the accompaniment as from the other (Butler, 1979b). This is true even with the orchestra arranged in nineteenth century fashion, with the first violins on one side and the second violins on the other side. Thus spatial separation by no means guarantees that music will be perceived in accor-

dance with the positioning of the instruments. Rather, groupings may be formed on the basis of some other attribute such as pitch, and this may in turn cause the listener to mislocalize the components of the musical configuration in accordance with such groupings. It also appears that other attributes such as loudness and timbre may be perceptually reorganized in this fashion.

Such findings, apart from their musical relevance, are of general interest to perceptual psychology, since they show that subjective grouping is not simply a matter of linking different stimuli together. Rather, this may involve a process in which the different stimulus attributes are dissociated and recombined so that illusory percepts result.

The experiments just described involved two musical sequences that were simultaneous or near-simultaneous. What happens when temporal differences are introduced? To examine this issue, Deutsch (1979) presented listeners with two melodic patterns, and they identified on each trial which one they had heard. Four conditions were employed. In the first, the melody was presented simultaneously to both ears, and here the level of identification performance was very high. In the second condition, the component tones of the melody switched between the ears, and here identification performance was considerably poorer. Subjectively in this condition the listener felt impelled to attend to the signal arriving at one ear or the other, and could not integrate the two sets of signals into a single perceptual stream. In the third condition, the component tones of the melody still switched between the ears; however the melody was accompanied by a drone. Whenever a component of the melody was in the right ear the drone was in the left ear, and whenever a component of the melody was in the left ear the drone was in the right ear. Thus the two ears again received input simultaneously, even though the melody to be identified still switched between the ears. This simultaneity of input produced a dramatic rise in identification performance. In the fourth condition, a drone was again presented; but this time to the same ear as the ear receiving the melody component (rather than the contralateral ear). Thus input was again to only one ear at a time. Here identification performance was again very low.

This experiment demonstrates that for tones emanating from different spatial locations, temporal relationships between them are important determinants of grouping. When signals are delivered to both ears simultaneously, it is easy to integrate the information into a single perceptual stream. But when the signals delivered to the two ears are clearly separated in time, subjective grouping by spatial location is so powerful as to prevent the listener from combining the signals to produce an integrated percept.

This finding leads one to ask what happens in the intermediate case, where the tones arriving at the two ears are not simultaneous, but rather overlapping in time. In a further experiment this intermediate case was found to produce intermediate results. Identification of the melody in the presence of the contralateral drone was poorer when the melody and drone were asynchronous than when they were strictly synchronous, but better than when there was no accompanying drone (Deutsch, 1979).

We can conclude from these studies that when a rapid sequence of tones is distributed between spatially separated instruments, and a clear temporal separation exists between the sounds produced by these instruments, the listener may be unable to integrate the sequence into a single coherent stream. However, a certain amount of overlap among the different instruments will facilitate such integration. Yet there is a tradeoff:

the greater the amount of overlap, the greater will be the loss of spatial distinctiveness; and as simultaneity is approached, spatial illusions may occur.

We now turn to the question of how perception of two simultaneous sequences of tones may be affected by whether the higher tones are presented to the right and the lower tones to the left, or whether this configuration is reversed. We noted earlier that, in the scale illusion, righthanders tend to perceive higher tones as on the right and lower tones as on the left, regardless of their actual locations. Thus simultaneous tone pairs of the "high-right/low-left" type tend to be well localized, and pairs of the "high-left/low-right" type tend to be mislocalized. This finding has been confirmed in more general settings (Deutsch, 1983).

We may then enquire whether pitch perception might also be affected by such spatial considerations. In an experiment to investigate this question, musically trained listeners were asked to notate two sequences of tones which were simultaneously presented, one to each ear. Tone pairs of which the higher was on the right and the lower on the left were notated significantly more accurately than tone pairs of which the higher was on the left and the lower on the right. This was found true with sequences organized in several different ways (Deutsch, 1983).

The above findings explain certain patterns of ear advantage which have been obtained for musical materials, and which have been thought to reflect patterns of hemispheric asymmetry in processing such materials. In addition, they have implications for the question of optimal seating arrangements for orchestras. In general, contemporary arrangements are such that, from the performers' point of view, instruments with high registers tend to be to the right, and instruments with low registers to the left. Figure 4 shows, for example, a seating arrangement of the Chicago symphony orchestra. From the above findings we can assume that this "high-right/low-left" disposition has evolved by trial and error because it is conducive to optimal performance. However, this leaves us with a paradox: From the viewpoint of the audience this configuration is mirror-imaged reversed, and so is such as to cause perceptual difficulties. There is no easy solution to this paradox for the case of concert hall listening (see Deutsch, in press, for a
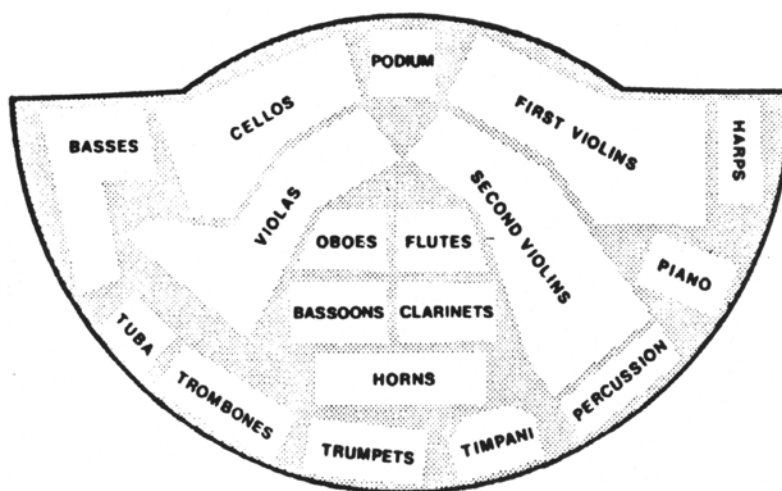

FIG. 4. Chicago Symphony seating plan from the viewpoint of the orchestra. (Adapted from Machlis, 1977.)

discussion). However, we may assume that reversing this disposition in multitrack recording should result in enhanced perceptual clarity.

Another issue concerning music and performing space involves the aesthetic effects of different auditory environments. As implied in Berlioz's statement "There is no such thing as music in the open air," the enclosed space of the concert hall contributes much to the aesthetic quality of music, through the complex sound reflections that are produced in this environment. The phenomenological effects of these reflections have frequently been discussed by musicians at an informal level, and recently they have been the subject of controlled experimental investigation. The physicist Schroeder and his associates have conducted a series of studies in which recordings of music were made in numerous European concert halls by means of two microphones placed at the ears of a "dummy." These recordings were then played to listeners in an anechoic chamber, enabling a realistic recreation of the acoustics of the concert hall at the ears of the "dummy." The method of paired comparisons was used to obtain preference ratings, and the individual scores were subjected to multidimensional scaling, thus producing a "preference space." Analyses of the correlations between various physical parameters of a concert hall and its coordinates in this "preference space" led to the conclusion that the greater the similarity of the signals arriving at the two ears, the lower the preference. This conclusion was reinforced by further studies in which the recorded signals were modified so as to increase binaural dissimilarities by adding lateral reflections. This manipulation had the expected effect of increasing preference ratings. It was concluded that wide halls with low ceilings (which tend to be constructed today for economic considerations) are associated with less listener enjoyment than narrow halls with high ceilings (more typical of older concert halls); since the latter type of design emphasizes early lateral reflections (Schroeder, 1980).

The study of spatial aspects of music is another area where the concerns of composers and of scientists have combined to very useful effect. Apart from their relevance to music, experiments on the effects of spatial separation have served to elucidate the nature of fundamental mechanisms involving stimulus integration and separation.

The Law of Stepwise Progression and the Principle of Proximity

In textbooks on tonal music we generally encounter the "law of stepwise progression, " which states that melodic progression should be by steps (a half step or a whole step) rather than by skips (more than a whole step), since stepwise progression is considered "stronger" or "more binding." What is left unspecified is why this law should be obeyed: The reader is supposed either to accept the law uncritically or to recognize its truth in some way by introspection.

To the psychologist, this law appears as an example of the Gestalt principle of proximity, which states that we tend to group together elements that are proximal along some dimension and to separate those that are spaced further apart (Wertheimer, 1923). Presumably, we have evolved mechanisms that produce such perceptual groupings, since this is conducive to an effective interpretation of our environment. Thus in the case of vision, proximal elements are more likely to belong together than elements that are

spaced further apart. In the case of hearing, sounds that are similar in frequency spectrum are likely to emanate from the same source, and sounds that are dissimilar are likely to be coming from different sources.

Consideration of the "law of stepwise progression" therefore leads us to enquire specifically into the ways in which the principle of proximity manifests itself when applied to pitch. Not only is this question of interest to perceptual psychology, but such enquiry also serves to provide the "law of stepwise progression" with a rational basis, by demonstrating the adverse effects to be expected when it is violated. Further, by characterizing the ways in which such effects behave under parametric manipulation, we can determine the conditions under which the law may be violated with relative impunity, and those under which its violation produces strongly adverse effects on perception and memory.

In an experimental setting, the impression of connectedness produced by a sequence of tones depends in a complex fashion on the pitch relationships involved, and also on their interaction with other factors (Deutsch, 1982a). One such factor is tempo. The higher the rate of presentation, the greater is the tendency for tones that are disparate in pitch to be heard as separate rather than as single connected series (Schouten, 1962). A second factor is attentional set. When presented with a sequence of two alternating tones, the listener may attempt to hear these either as a single connected series or as two disjoint series. As shown in Figure 5, when the listener is attempting to hear a single series, the impression of connectedness depends very strongly on presentation rate. However, when the listener is attempting to hear the tones as disconnected, temporal factors appear unimportant (Van Noorden, 1975). A third factor is the length of sequence presented. For an impression of connectedness to be obtained, a larger decrease in tempo is required for long sequences than for two-tone sequences (Van Noorden, 1975).

One adverse effect of violating the principle of proximity, at least at fast tempi, is that temporal relationships between adjacent tones become difficult to judge. For example, when a rapid sequence of tones is presented, and these are drawn from two different
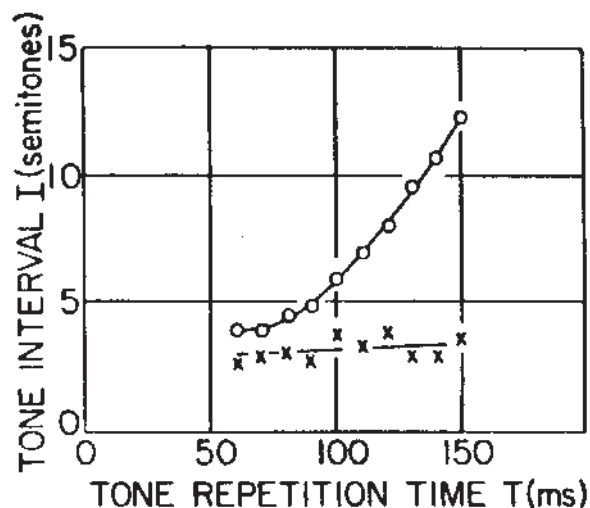

FIG. 5. Boundaries for perception of a sequence of tones as a connected series as a function of pitch proximity and tempo. (o) Listener attempting to hear a a connected series. (x) Listener attempting to hear a disconnected series. (From Van Noorden, 1975.)

pitch ranges, judgment of the orders of these tones is very difficult. However this problem disappears when the tones are brought close together in pitch (Bregman & Campbell, 1971). When the presentation rate is slowed down, so that order perception is readily accomplished, there is still a gradual breakdown of temporal resolution as the pitch disparity in a sequence of alternating tones increases. For example, it becomes increasingly difficult to detect a rhythmic irregularity in such a sequence. This effect is also more pronounced with long sequences than with short ones (Van Noorden, 1975).

A further loss of perceptual accuracy that results from violating the principle of proximity involves the situation where two simultaneous sequences of tones are presented, each in a different spatial location. As described earlier, there is a tendency to reorganize such sequences perceptually in accordance with pitch proximity, so that a sequence formed by tones in one pitch range appears to be coming from one spatial location and a sequence formed by tones in a different pitch range appears to be coming from the other location (Butler, 1979a; Deutsch, 1975a, 1975b, 1979). This phenomenon is also related to another musical rule which forbids the crossing of voices in counterpoint. If the composer attempts to produce a crossing of voices, there is a risk that the listener will synthesize voices in accordance with pitch proximity rather than in accordance with the composer's intentions. This perceptual phenomenon holds true also when only a single spatial location is involved (though the illusion that tones in one pitch range are emanating from one spatial location and tones in another pitch range from a different location is of course not produced).

Finally, pitch proximity can be shown to affect the ability to recognize individual tones in a sequence. Deutsch (1978a) employed the following paradigm. Listeners compared the pitches of two tones when these were separated by a time interval during which a sequence of extra tones was interpolated. They were asked to ignore the interpolated tones and to judge whether the test tones were the same or different in pitch. Accuracy of pitch recognition was found to increase as the average size of the intervals formed by the interpolated tones decreased. It was concluded that the interpolated sequence provides a framework of pitch relationships in which the test tones are embedded and that the more proximal these relationships the stronger the framework.

The perceptual separation that occurs between tones that are disparate in pitch can be exploited to musical advantage. If a composer wishes the listener to perceive two simultaneous melodic lines, this can be greatly facilitated by presenting the two lines in different pitch ranges. A particularly interesting technique that exploits this phenomenon was used extensively by the Baroque composers and is known as pseudopolyphony. Here an instrument plays a rapid sequence of single tones which are drawn from two different pitch ranges; as a result the listener perceives two melodic lines in parallel. Dowling (1973) has demonstrated the strength of this perceptual effect in a formal experiment. He presented listeners with two well-known melodies, which were interleaved in time. The listeners were asked to identify the melodies. When these were drawn from the identical pitch range the task was very difficult, since temporally adjacent tones were perceptually combined into a single stream. However, as one of the interleaved melodies was gradually transposed, so that the pitch ranges of the two melodies diverged, identification became increasingly more easy.

The above studies demonstrate the usefulness of the experimental technique in understanding the basis of musical rules which have developed by trial and error. The conclusions from these studies could not have been arrived at by examination of musical examples alone, and many of them are not apparent from introspection.

Musical Shape Analysis and the Theory of Twelve-Tone Composition

Present-day interest in shape analyzing mechanisms has stemmed largely from the work of the Gestalt psychologists at the end of the last century and the beginning of this one. The Gestaltists were concerned with characterizing the ways in which shapes may be transformed without losing their perceptual identities. For example, the identities of visual shapes are not destroyed when they are changed in size or translated to a different position in the visual field (Sutherland, 1973).

The large majority of work on shape analysis has been concerned with vision. However, it may be noted that Von Ehrenfels (1890) in his influential paper "Uber Gestltqualitaten" gave melody as an example of a Gestalt. He pointed out that a melody when transposed retains its essential form, the *Gestaltqualitat*, provided that the relations among individual tones are unaltered. In this respect, he argued, melodies are like visual shapes.

Largely unknown to psychologists, the theory of twelve-tone composition, developed early in this century by Schoenberg, is based on a theory of shape analysis for pitch structures. This theory is in turn based on an intermodal analogy in which one dimension of visual space is mapped into pitch and another into time. Describing his system of composition as "not a mere technical device" but as of the "rank and importance of a scientific theory," Schoenberg justifies it in the following way:

> THE TWO-OR-MORE DIMENSIONAL SPACE IIN WHICH MUSICAL IDEAS ARE PRESENTED IS A UNIT. . . . The elements of a musical idea are partly incorporated in the horizontal plane as successive sounds, and partly in the vertical plane as simultaneous sounds. . . . *The unity of musical space demands on absolute and unitary perception*. In this space. . .there is no absolute down, no right or left, forward or backward. . . . To the imaginative and creative faculty, relations in the material sphere are as independent from directions or planes as material objects are, in their sphere, to our perceptive faculties. Just as our mind always recognizes, for instance, a knife, a bottle or a watch, regardless of its position, and can reproduce it in the imagination in every possible position, even so a musical creator's mind can operate subconsciously with a row of tones, regardless of their direction, regardless of the way in which a mirror might show the mutual relations, which remain a given quantity [Schoenberg, 1951, pp. 220-223].

Figure 6 illustrates Schoenberg's use of his theory in compositional practice. As he wrote: "The employment of these mirror forms corresponds to the principle of *the absolute and unitary perception of musical space* [p. 225]."

Schoenberg thus proposed that a tone row, defined as a particular linear ordering of the twelve tones of the chromatic scale, retains its perceptual identity under the following transformations: when it is transposed to a different pitch range ("transposition"), when all ascending intervals become descending intervals and vice versa ("inversion"), when it is presented in reverse order ("retrogression"), and when it is transferred by both
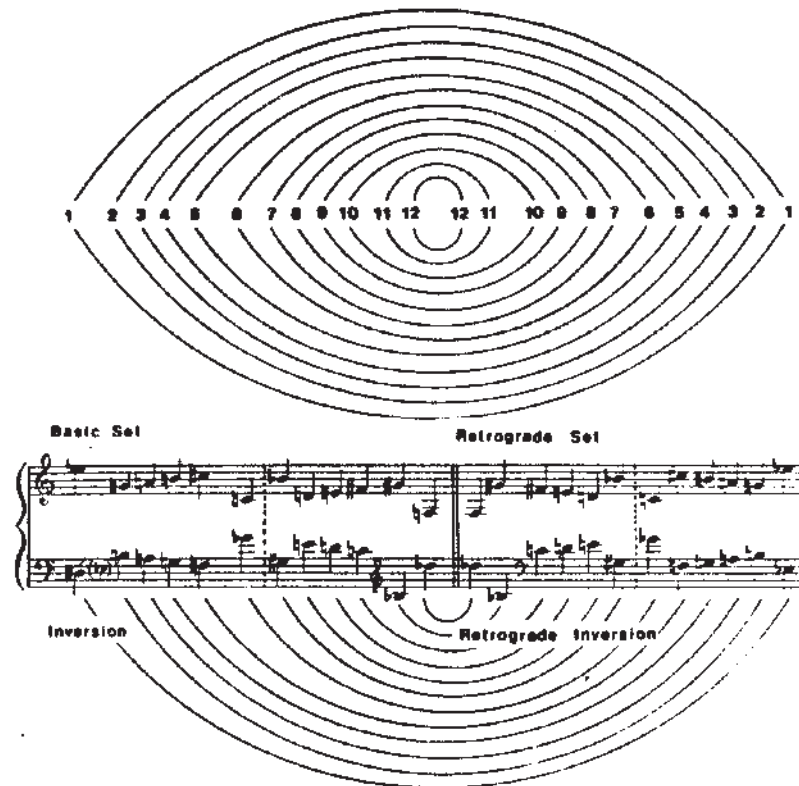
FIG. 6. Schoenberg's illustration of his theory of equivalence relations between pitch structures. The
musical example is taken from the Wind Quartet, Op. 26. (From Schoenberg, 1951.)

these operations ("retrograde-inversion"). Further, Schoenberg proposed that, given the strong perceptual similarity between tones than are separated by octaves, the identity of a tone row is preserved when the individual tones in the row are placed in different octaves.

Schoenberg's theory provided the basis for much sophisticated system building around the middle of the century. Foremost here is the work of Babbitt and his followers in interpreting the twelve-tone system as a group. The elements of the group are twelve-tone sets, represented as permutations of pitch or order numbers; the operation is the multiplication of permutations (Babbitt, 1960, 1961). This system has been used extensively in compositional practice (see also Perle, 1972, 1977).

The question may be raised of whether the equivalence relations defined in twelve-tone theory are indeed utilized by the perceptual system. We may note that Schoenberg's intermodal analogy, although interesting, is rather forced. It makes sense to assume that we have evolved mechanisms that enable us to recognize an object when it is presented in a different orientation relative to the observer. However, it does not make sense in the same way to assume that we will recognize a sound sequence when it is reversed in time or when its pitch relationships are turned upside-down: In our natural environment we are never required to do this. Further, it has been shown in the case of vision that some formal relationships that exist within a configuration are readily perceived, others are perceived with difficulty, and yet others are not perceived at all (Garner, 1974).

Concerning the perceptual identity of a tone row under retrogression and inversion,

two studies in the psychological literature may be cited. White (1960) used a long-term recognition paradigm to study the ability of listeners to identify well-known melodies when these were played in retrogression. Some recognition was obtained; however, performance was no better than when the melody was played in a monotone with rhythm as the only cue. Further, better recognition was obtained when the intervals within the melody were randomly permuted than when the orders of the tones were strictly reversed. This indicates that the listeners were recognizing the retrograde sequences on the basis of the set of intervals involved, rather than on their orderings.

Dowling (1972) used a short-term paradigm to study recognition of a sequence of tones under retrogression, inversion, and retrograde-inversion. He presented listeners with a standard five-tone sequence, followed by a comparison sequence. The comparison was either unrelated to the standard, or it was an exact transposition, or it was transformed by retrogression, inversion, or retrograde-inversion. In another set of conditions, the comparison sequence was further distorted so that its contour was preserved but the exact intervals were not. Although the listeners performed above chance on these tasks, they were unable to distinguish between exact transformations and those that preserved contour alone. Dowling (1978) later provided evidence that exact interval recognition was being masked by the listeners' projecting the pitch information onto the highly over-learned scales of our tonal system. Whether extensive exposure to twelve-tone music could overcome such a masking effect is a matter that requires further investigation.

Another issue raised by twelve-tone theory is whether a sequence of tones retains its perceptual identity when its components are placed in different octaves. For   single tones in isolation, there is a strong perceptual similarity between tones that stand in octave relation. Psychologists have noted this equivalence and refer to tones that are an octave apart as having the same "tone chroma" (Bachem, 1954; Meyer, 1904, 1914; Revesz, 1913; Ruckmick, 1929; Shepard, 1964; Ward & Burns, 1982). Further, traditional music theory recognizes the equivalence of such tones in simultaneous structures through the rules governing chord progressions (Rameau, 1722/1950). However, where melodies or successive pitch structures are concerned, octave equivalence does not obviously hold, since we do not interchange octaves in successive contexts in the same way as we do in simultaneous contexts.

According to twelve-tone theory, tones that are separated by octaves are considered to be in the same "pitch class," and their equivalence is assumed to be a perceptual invariant. It is therefore held that intervals (both simultaneous and successive) retain their perceptual identities when the tones forming these intervals are placed in different octaves; such intervals are held to be in the same "interval class." However, the hypothesized equivalence relation of interval class is not a necessary consequence of interval equivalence together with octave equivalence. Deutsch (1969) proposed a neural network for the abstraction of pitch combinations in which the perceptual equivalence of transposed intervals and chords is mediated by one channel, and the perceptual equivalence of tones that are separated by octaves, together with the invertibility of chords, is mediated by a separate and parallel channel. This network gives rise to octave equivalence for single tones in isolation and in a harmonic or simultaneous context, but not in a melodic or successive context.

In an experiment designed to examine the issue of octave equivalence in a successive

context, the tune "Yankee Doodle" was presented to listeners in several versions (Deutsch, 1972). One version was untransformed. In a second version, the tones were in their correct positions within the octave, but the octaves in which they were placed varied randomly; thus interval class was preserved even though the intervals were altered. In a third version, the pitch information was removed entirely. Each version was played to a different group of listeners. Although the untransformed version was recognized by everyone, recognition of the randomized octaves version was no better than of the version where the pitch information was removed entirely. This finding is as predicted from the two-channel model of Deutsch (1969), and it shows that interval class cannot be treated as a perceptual invariant.

When listeners in this study were later informed of the identity of the melody and heard it again, many found that they could now follow it to a large extent and confirm that each note was indeed correctly placed within its octave. Thus the listeners were able to use octave equivalence to confirm a hypothesized melodic shape, though they were unable to recognize this shape in the absence of strong cues on which the hypothesis might be based. We can conclude that interval class can be perceived in a successive context under certain conditions, but that such perception does not result from a passive process. Rather, it may be regarded as an example of "top-down" shape analysis; i.e., as the result of hypothesis-testing by the listener.

Further studies support this argument. Dowling and Hollombe (1977) presented listeners with melodies whose individual tones were placed in different octaves, and they found that recognition performance was better for melodies whose contours were preserved than for melodies with altered contours. This finding is in accordance with the present line of reasoning. Since melodies can be recognized on the basis of their contours alone (Werner, 1925; White, 1960), contour should act as a powerful cue for hypothesis-testing. Similar findings were obtained by Idson and Massaro (1978) and Kalman and Massaro (1979). Second, it has been found that when listeners were presented with a small set of melodies many times and were asked to identify each melody from a small list of alternatives, recognition performance was considerably better than when such melodies were presented only once with no cues concerning their identity (House, 1977; Idson & Massaro, 1978). In the former case ample opportunity was given for hypothesis testing, so that again enhanced recognition would be expected (Deutsch, 1978b).

Returning to twelve-tone theory, we can conclude that interval class may be perceived, but only under conditions of reasonably high expectancy. The ability of a listener to recognize a tone row under octave displacement should depend critically on such factors as prior familiarity with the row and whether or not the relationships formed by earlier tones in the row are such as to produce clear expectations for the later tones (see also Deutsch, in 1982b).

Hierarchical Structure in Music

It may generally be stated that we tend to encode and retain information in the form of hierarchies when given the opportunity to do so. For example, programs of behavior tend to be retained as hierarchies (Miller, Galanter, & Pribram, 1960) and goals in prob-

lem solving as hierarchies of subgoals (Ernst & Newell, 1969). Visual scenes appear to be encoded as hierarchies of subscenes (Hanson & Riseman, 1978; Navon, 1977; Palmer, 1977; Winston, 1973). The phase structure of a sentence lends itself readily to hierarchical interpretations (Chomsky, 1963; Johnson-Laird, in this volume; Miller & Chomsky, 1963; Yngve, 1960). When presented with artificial serial patterns which may be hierarchically encoded, we readily form encodings that reflect pattern structure (Kotovsky & Simon, 1973; Restle, 1970; Restle & Brown, 1970; Simon & Kotovsky, 1963; Vitz & Todd, 1967, 1969). Such findings have given rise to the development of sophisticated models of serial pattern representation in terms of heirarchies of operators (Greeno and Simon, 1974; Leewenberg, 1971; Restle, 1970; Simon, 1972; Simon and Kotovsky, 1963; Simon and Sumner, 1968; Vitz and Todd, 1967, 1979).

In considering how we most naturally form hierarchies, however, theories have generally been constrained by the nature of the stimulus material under consideration. For example, visually perceived objects are naturally formed out of parts and subparts. The hierarchical structure of language must necessarily be constrained by the logical structure of events in the world. The attainment of a goal is generally arrived at by an optimal system of subgoals. And so on.

This problem is just as severe for theories based on experiments utilizing artificial serial patterns devised by the experimenter. To take a concrete example, Restle's (1970) theory of hierarchical representation of serial patterns evolved from findings based on the following experimental paradigm. Subjects were presented with a row of six lights, which turned on and off in repetitive sequence, and they were required on each trial to predict which light would come on next. The sequences were structured as hierarchies of operators. For example, given the basic subsequence X = (1 2), then the operation R ('repeat of X') produces the sequence 1 2 1 2; the operation M ('mirror-image of X') produces the sequence 1 2 6 5, and the operation T ('transposition +1 of X') produces the sequence 1 2 2 3. Through recursive application of such operations, long sequences can be generated which have compact structural descriptions. Thus M(T(R(T(1)))) describes the sequence 1 2 1 2 2 3 2 3 6 5 6 5 5 4 5 4.

Restle and Brown (1970), using sequences constructed in this fashion, found compelling evidence that subjects encoded them in accordance with their hierarchical structure. However, it should be noted that the sequences were constructed so as to allow for only one hierarchical interpretation. Thus it is difficult to estimate the generalizability of this model to situations where alternative hierarchical realizations are possible.

Given these problems, the hierarchical structure of established music is of particular interest to cognitive psychology, since such music is solely the product of human processing mechanisms, unfettered by external constraints. Further, music can reasonably be considered to have evolved so as to make optimal use of these mechanisms.

Long before cognitive psychologists became seriously interested in hierachical structure, the music theorist Schenker proposed a hierarchical system for tonal music that has points of similarity with the system proposed by Chomsky for language (Chomsky, 1957, 1965). (In fact, Schenker acknowledged that his ideas were inspired by C. P. E. Bach whose *Essay on the True Art of Playing Keyboard Instruments* details the processes by which a simple musical event may be replaced by a more elaborate musical event which expresses the same basic content.) In Schenker's system, music is regarded as a hierar-

chy in which notes at any given level are considered "prolonged" by a sequence of notes at the next-lower level. Three basic levels are distinguished. First there is the *foreground*, or surface representation; second there is the *middleground*; and third there is the *background*, or *Ursatz*. The *Ursatz* is itself considered a prolongation of the triad (Schenker, 1956, 1973).

Schenker's work, though largely unrecognized in his time, has had a profound influence on music theory since the late 1950s (see, e.g., Forte, 1974; Salzer, 1962; Westergaard, 1975; Yeston, 1977). Most Schenkerian analysis, however, is purely descriptive in nature and is generally regarded as an end in itself. Furthermore, the assumptions of Schenkerian analysis are at basis rather inexplicit.

The collaborative work of the music theorist Lerdahl and the psycholinguist Jackendoff (1977) represents an attempt to explicate the structure of Schenker's system and to interpret this structure as a form of internal representation. Their approach makes use of tree diagrams that resemble in some respects those used in transformational grammar. However, the authors are careful to emphasize the very real differences that exist between language and music. For example, linguistic trees represent "is-a" relations: A noun phrase that is followed by a verb phrase *is a* sentence, and so on. In contrast, musical trees do not involve grammatical categories. Rather, the fundamental relationship that they express is that of the elaboration of a single pitch event by a sequence of pitch events. Their theory also emphasizes the importance of psychological grouping phenomena in the formation of musical hierarchies.

Schenker's theory is essentially "top-down" in nature, in that the *Ursatz* acts as a "kernel" from which the middleground and foreground structures are derived. (This is analogous to transformational grammar, which relies on "kernel" sentences to generate linguistic structures.) The foreground levels are held to be generated from above, from levels at which the actual notes are not themselves present. The music theorist Narmour (1977) has argued that this constitutes a serious difficulty for Schenker's theory. He shows by numerous examples that patterns of relationship between notes that are not necessarily adjacent at the foreground level contribute importantly to musical structure. He proposes alternatively that a given representation is generated "bottom-up" and that Schenker's terminal symbols (the actual notes on the page of the composition) be conceived not as the result of mappings onto a lower level from middleground structures and the background kernel, but rather as the initiating structure from which higher-level structures are built. He also argues that foreground structures create multiple alternative representations (or implications), so that musical pieces should be conceptualized not as tree structures, but rather as interlocking networks.

Narmour's work was inspired by that of the music theorist Meyer (1956, 1960, 1973), who argues that musical structure should be viewed in terms of implications generated by pitch events that are realized by further pitch events. Such implications and their realizations are considered to occur at all hierarchical levels. Further, a sequence of pitch events often has multiple implications, only some of which are realized.

Deutsch and Feroe (1981) have advanced a formal theory for the representation of pitch sequences in tonal music which falls into the class of those developed by Restle, Simon, and others in that it proposes a specific language or notation for describing serial patterns, and this language is considered to reflect specific encodings. However, the

concerns of music theorists were also considered in developing the formalism. Basically pitch sequences are assumed to be retained as hierarchical networks. Elements that are present at each hierarchical level are elaborated by further elements at the next-lower level, until the lowest level is reached. At each level of the hierarchy, elements are organized as structural units in accordance with laws of figural goodness. The basic architecture of this system can also be applied to the internal representation of other types of information, such as visual scenes (Lynch, 1960).

Hierarchical structure in music provides a rich field for experimental investigation, which has so far been largely untapped. Two recent studies may be mentioned. The psychologist Rosner in collaboration with the music theorist Meyer (1982) addressed the following question: Frequently, melodies appear to be hierarchically structured in such a way that the type of patterning exhibited by a given melody changes from one hierarchical level to the next. For example, a melody at one level may be characterized by a linear pattern; at another level by a gap-fill pattern;[5] and so on. The authors further hypothesized that melodies are classified by the listener in terms of the organization at the highest level at which significant closure is created. In a test of this hypothesis, musically untrained listeners were asked to categorize melodies in a concept identification task. The melodies had previously been categorized by musical analysis as either gap-fill or non-gap-fill at the appropriate hierarchical level. It was found that listeners classified the melodies in accordance with theoretical expectations.

In another study, Deutsch (1980) compared memory for tonal sequences that were hierarchically structured with those that were not. Musically trained listeners were presented with sequences which they recalled in musical notation. Half of the sequences were hierarchically structured such that a higher-level subsequence of three elements acted on a lower-level subsequence of four elements. The remaining sequences were unstructured. Recall was found to be considerably superior for the structured than for the unstructured sequences. It was concluded that listeners readily detect hierarchical organization in tonal sequences and can utilize this organization so as to produce parsimonious encodings.

In summary, hierarchical structure in music is an area of study which has strong dividends for both music theory and psychology. Since music is the product of our processing mechanisms and since traditional music may be taken to have evolved so as to make optimal use of these mechanisms, understanding the structure of tonal music and how it is processed is likely to have broad implications for theories of human cognition.

CONCLUSIONS: MUSIC THEORY AND PSYCHOLOGY

In the Introduction I referred to fundamental problems in determining the relationship of findings in psychology to music theory. It is with a discussion of these problems that I shall conclude this essay.

Psychology contributes to the understanding of music by characterizing the processing mechanisms of the listener. What is worrisome to some music theorists is the possibility that findings from psychology might be taken as a basis for arguing what music

---

[5] A gap-fill pattern is characterized by two elements: (a) a skip or succession of skips that move in the same direction and (b) a succession of steps which fill the gap, that move in the opposite direction.

ought to be. Much work in perceptual and cognitive psychology has to do with determining limits: limits to the amount of information that can be retained, limits of discriminability, and so on. Taking such "scientifically established limits" too seriously, it is feared, might serve to stultify musical development by creating artificial boundary conditions for acceptable music. For the limitations determined by such experiments might not in fact be fixed but might rather be a function of the type of music to which the listener has been exposed.

To place this concern in historical perspective, the development of Western music may be viewed as a constant struggle between innovative composers on the one hand and establishment critics on the other, who have argued against various innovations on the grounds that they are unacceptable to the listener. Some examples of "new" music that were considered unacceptable would surprise a modern audience. For example, J. S. Bach was considered in his time to have "confused the congregation with many peculiar and foreign tunes [Portnoy, 1954, p. 144]." Another composer who was censured by his contemporaries was Monteverdi. The distinguished music critic and theorist Artusi wrote of his music:

> Insofar as it introduced new rules, new modes, and new turns of phrase, these were harsh and little pleasing to the ear, nor could they be otherwise, for as long as they violate the good rules-in part founded by experience, the mother of all things, in part observed by nature, and in part by demonstration-we must believe them to be deformations of the nature and propriety of true harmony, far removed from the object of music [Artusi, 1600/1950. p. 394].

Yet the works of Bach and of Monteverdi appear to us as outstanding examples of traditional cultivated music. Clearly, the way that music affects the listener is at least to some extent a function of experience.

It should be stated that in the past, arguments against new music have been aesthetic in nature and were not based on controlled experiments demonstrating processing limitations. The possibility remains, however, that the typical listener of Monteverdi's time might have displayed a different set of processing limitations than those displayed by the typical listener of our time. One could plausibly regard the development of Western music as in part an extensive long-term field study in which generations of audiences have been exposed to various types of music and their processing mechanisms have been shaped and reshaped as a result of such exposure. It is this line of reasoning that causes some theorists to insist that when laboratory studies show that listeners do not perceive equivalences that exist formally in a musical system, this provides no argument against the ultimate viability of the system.

However, to dismiss the findings of psychology because of such concerns does not constitute a solution. If a music theory is to be scientifically justified, such justification must lie in its relationship to the processing mechanisms of the listener. To take an extreme example, no one would seriously consider composing in a musical system that employs only sounds outside the range of hearing. Central processing limitations are no less real than those of our peripheral hearing apparatus; the only difference is that while peripheral limitations are fixed, some central limitations are fixed and some are plastic.

There remains the question of determining which of our musical processing mecha-

nisms can be shaped by experience. To me it appears that no clear answer can be obtained by laboratory experimentation. We can expose subjects to intensive training on a given system and determine whether or not they can learn to use its rules. But negative results would not be conclusive, since it could always be argued that long-term exposure, particularly during early childhood, might have produced positive results instead. We can, however, make some inspired guesses as to which processing characteristics are likely to be fixed. Those characteristics which are most useful in making sense of our auditory environment are prime candidates. These include the tendency to fuse together components of a sound spectrum that are in harmonic relationship; the tendency to form sequential configurations on the basis of frequency proximity; the tendency to attend on the basis of spatial location; and so on. Such mechanisms are likely either to be hardwired or, if acquired through experience, to continue to be acquired as a result of experience with our nonmusical auditory environment. Amongst other candidates for fixed processing characteristics are those that lead to parsimony of encoding and other measures of encoding efficiency.

To conclude, it must remain the prerogative of the composer to experiment with any new rules that he wishes; psychology cannot provide prescriptive answers and can only explain how existing music is perceived. However, by the same token, music theory cannot provide prescriptive answers either. As Aristoxenus (1902, p. 195) wrote over two millennia ago: "We shall advance to our conclusions by strict demonstration." If there is no strict demonstration, then there can be no conclusions.

ACKNOWLEDGMENT

REFERENCES

Aristotle. [*De caelo*] (J. L. Stocks, trans.). In *The works of Aristotle* (Vol. 2). Oxford: Oxford University Press, 1930.

Aristoxenus. [*The harmonics of Aristoxenus*] (H. S. Macran, trans.). Oxford: Clarendon Press, 1902.

Artusi. *L'Artusi, ovvero, Delle imperferzioni della moderna musica*. In O. Strunk (Ed.), *Source readings in music history*. New York: Norton, 1950. (Originally published, 1600.)

Babbitt, M. Twelve-tone invariants as compositional determinants. *The Musical Quarterly*, 1960, *46*, 246-259.

Babbitt, M. Set structure as a compositional determinant. *Journal of Music Theory*, 1961, 5, 73-94.

Bach, C. P. E. [*Essay on the true art of playing keyboard instruments*] (W. J. Mitchell, Ed. and trans.). New York: W. W. Norton, 1949.

Bachem, A. Time factors in relative and absolute pitch determination. *Journal of the Acoustical Society of America*, 1954, *26*, 751-753.

Berger, K. W. Some factors in the recognition of timbre. *Journal of the Acoustical Society of America*, 1964, *36*, 1888-1891.

Berlioz, H. [*Treatise on instrumentation*] (R. Strauss, Ed. & T. Front, trans.). New York: E. F. Kalmus, 1948.

Boethius. [*Boethius' the principles of music*] (C. M. Bower, trans.). Ann Arbor, MI: University of Michigan Press, 1967.

Brant, H. Space as an essential aspect of musical composition. In E. Schwartz & B. Childs (Eds.),

Contemporary composers on contemporary music. New York: Holt, Rinehart and Winston, 1966.

Bregman, A. S. The formation of auditory streams. In J. Requin (Ed.), *Attention and performance VII*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1978.

Bregman, A. S., & Campbell, Jr. Primary auditory stream segregation and perception of order in rapid sequence of tones. *Journal of Experimental Psychology*, 1971, *89*, 244-249.

Bregman, A. S., & Pinker, S. Auditory streaming and the building of timbre. *Canadian Journal of Psychology*, 1978, *32*, 20-31.

Butler, S. *Hudibras, parts I and II, and selected other writings* (J. Wilders & H. de Quehen, Eds.). Oxford: Clarendon Press, 1973.

Butler, D. A further study of melodic channeling. *Perception and Psychophysics*, 1979, *25*, 264-268. (a)

Butler, D. Melodic channeling in a musical environment. *Research Symposium on the Psychology and Acoustics of Music*, Kansas, 1979. (b)

Chomsky, N. *Syntactic structures*. The Hague: Mouton, 1957.

Chomsky, N. Formal properties of grammars. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 2). New York: Wiley, 1963.

Chomsky, N. *Aspects of the theory of syntax*. Cambridge, MA: M.I.T. Press, 1965.

de Boer, E. On the "residue" and auditory pitch perception. In W.D. Keidel & W. D. Neff (Eds.), *Handbook of sensory physiology* (Vol. V/3). Wein: Springer-Verlag, 1976.

Deutsch, D. Music recognition. *Psychological Review,* 1969, *76*, 300-307.

Deutsch, D. Octave generalization and tune recognition. *Perception and Psychophysics*, 1972, *11*, 411-412.

Deutsch, D. Musical illusions. *Scientific American*, 1975, *233*, 92-104. (a)

Deutsch, D. Two-channel listening to musical scales. *Journal of the Acoustical Society of America*, 1975, *57*, 1156-1160. (b)

Deutsch, D. Delayed pitch comparisons and the principle of proximity. *Perception and Psychophysics*, 1978, *23*, 227-230. (a)

Deutsch, D. Octave generalization and melody identification. *Perception and Psychophysics*, 1978, 23, 91-92.(b)

Deutsch, D. Binaural integration of melodic patterns. *Perception and Psychophysics*, 1979, *25*, 399-405.

Deutsch, D. The processing of structured and unstructured tonal sequences. *Perception and Psychophysics*, 1980, *28*, 381-389.

Deutsch, D. Grouping mechanisms in music. In D. Deutsch (Ed.), *The psychology of Music*. New York: Academic Press, 1982. (a)

Deutsch, D. The processing of pitch combinations. In D. Deutsch (Ed.), *The psychology of music*. New York: Academic Press, 1982. (b)

Deutsch, D. Dichotic listening to musical sequences: Relationship to hemispheric specialization of function. *Journal of the Acoustical Society of America*, 1983, *74*, S79-80.

Deutsch, D. Musical space. In W. R. Crozier & A. J. Chapman (Eds.), *Cognitive processes in the perception of art*. Amsterdam: North Holland, in press.

Deutsch, D., & Feroe J. The internal representation of pitch sequences in tonal music. *Psychological Review*, 1981, *88*, 503-522.

Dowling, W. J. Recognition of melodic transformations: Inversion, retrograde, and retrograde-inversion. *Perception and Psychophysics*, 1972, *12*, 417-421.

Dowling, W. J. The perception of interleaved melodies. *Cognitive Psychology*, 1973, *5*, 322-377.

Dowling, W. J. Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, 1978, *85*, 342-354.

Dowling, W. J., & Hollombe, A. W. The perception of melodies distorted by splitting into several octaves: Effects of increasing proximity and melodic contour. *Perception and Psychophysics*, 1977, *21*, 60-64.

Ehrenfels, C. Von. Uber Gestaltqualitaten. *Vierteljahrschrift fur Wissenschaftliche Philosophie*, 1890, *14*, 249-292.

Erickson, R. *Sound structure in music*. Berkeley: University of California Press, 1975.

Erickson, R. New music and psychology. In D. Deutsch (Ed.), *The psychology of music*. New York: Academic Press, 1982.

Ernst, G. W., & Newell, A. *GPS: A case study in generality and problem solving*. New York: Academic Press, 1969.

Forte, A. *Tonal harmony in concept and practice*. New York: Holt, Rinehart and Winston, 1974.

Freeman, K. *Ancilla to the pre-Socratic philosophers*. Cambridge, MA: Harvard University Press, 1948.

Garner, W. R. *The processing of information and structure*, New York: Wiley, 1974.

Greeno, J. G., & Simon, H. A. Processes for sequence production. *Psychological Review*, 1974, *81*, 187-196.

Gregory, R.L. *The intelligent eye*. New York: McGraw-Hill, 1970.

Grey, J. M. *An exploration of musical timbre*. Unpublished doctoral dissertation. Stanford University, 1975.

Grey, J. M., & Moorer, J. A. Perceptual evaluation of synthesized musical instrument tones. *Journal of the Acoustical Society of America*, 1977, *62*, 454-462.

Hanson, A. R., & Riseman, E. M. (Eds.). *Computer vision systems*. New York: Academic Press, 1978.

Hawkins, Sir J. A. *General history of the science and practice of music* (Vol.1). London: Dover, 1963. (Originally published, 1853.)

Helmholtz, H. Von. [*Helmholtz's physiological optics*] (J. P. C. Southall, Ed. and trans.). Rochester, New York: Optical Society of America, 1925. (Originally published, 1909-1911.)

Helmholtz, H. von. *On the sensations of tone as a physiological basis for the theory of music*. New York: Dover, 1954. (Originally Published, 1885.)

Hirsh, I. J. Auditory perception of temporal order. *Journal of the Acoustical Society of America*, 1959, *31*, 759-767.

Hirsh, I. J., & Sherrick, C. E. Perceived order in different sense modalities. *Journal of Experimental Psychology*, 1961, *62*, 423-432.

Hochberg, J. Organization and the Gestalt tradition. In E.C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol.1). New York: Academic Press, 1974.

House. W. J. Octave generalization and the identification of distorted melodies. *Perception and Psychophysics*, 1977, *21*, 586-589.

Hunt, F. V. *Origins in acoustics*. New Haven, CT: Yale University Press, 1978.

Idson, W. L., & Massaro, D. W, A bidimensional model of pitch in the recognition of melodies. *Perception and Psychophysics*, 1978, *24*, 551-565.

Kallman, H. J., & Massaro, D. W. Tone chroma is functional in melody recognition. *Perception and Psychophysics*, 1979, *26*, 32-36.

Kotovsky, K., & Simon, H. A. Empirical tests of a theory of human acquisition of concepts for sequential events. *Cognitive Psychology*, 1973, *4*, 399-424.

Leewenberg, E. L. A perceptual coding language for visual and auditory patterns. *American Journal of Psychology*, 1971, *84*, 307-349.

Lerdahl, F., & Jackendoff, R. Toward a formal theory of music. *Journal of Music Theory*, 1977, *21*, 111-172.

Lynch, K. *The image of the city*. Cambridge, MA: Harvard University Press, 1960.

Machlis, J. *The enjoyment of music*. New York: Norton, 1977.

McAdams, S. Spectral fusion and the creation of auditory images. In M. Clynes(Ed.), *Music, mind and brain*. New York: Plenum Press, 1981.

Mathews, M. V. *The technology of computer music*. Cambridge. MA: M.I.T. Press, 1969.

Mathews, M.V., & Pierce, J. R. Harmony and nonharmonic partials. *Journal of the Acoustical Society of America*, 1980, *68*, 1252-1257.

Meyer, L.B. *Emotion and meaning in music*. Chicago, IL: University of Chicago Press, 1956.

Meyer, L.B. *Music, the arts and* ideas. Chicago, IL: University of Chicago Press, 1960.

Meyer, L.B. *Explaining music: Essays and explorations*. Berkeley, CA: University of California Press, 1973.

Meyer, M. On the attributes of the sensations. *Psychological Review*, 1904, *11*, 83-103.

Meyer, M. Review of G. Revesz, "Zur Grundleguncy der Tonpsychologie." *Psychological Bulletin*, 1914, *11*, 349-352.

Miller, G. A., & Chomsky, N. Finitary models of language users. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 2) New York: Wiley, 1963.

Miller, G. A., Galanter, E. H., & Pribram, K. H. *Plans and the structure of behavior*. New York: Holt, Rinehart and Winston, 1960.

Narmour, E. *Beyond Schenkerism*. Chicago: University of Chicago Press, 1977.

Navon, D. Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, 1977, *9*, 353-383.

Palisca, C.V. Scientific empiricism in musical thought. In H. H. Rhys (Ed.), *Seventeenth century science in*

the arts. Princeton, NJ: Princeton University Press, 1961.

Palmer, S.E. Hierarchical structure in perceptual representation. *Cognitive Psychology*, 1977, *9*, 441-474.

Perle, G. *Serial composition and atonality*. Berkeley, CA: University of California Press, 1972.

Perle, G. *Twelve-tone tonality*. Berkeley, CA: University of California Press, 1977.

Plomp, R. The ear as frequency analyzer. *Journal of the Acoustical Society of America*, 1964, *36*, 1628-1636.

Plomp, R. Timbre as a multidimensional attribute of complex tones. In R. Plomp & G.F. Smoorenburg (Eds.), *Frequency analysis and periodicity detection in hearing*. Sijthoff: Leiden, 1970.

Plomp. R., & Mimpen, A. M. The ear as frequency analyzer II. *Journal of the Acoustical Society of America*, 1968, *43*, 764-767.

Plomp, R., & Steeneken, H. J. M. *Pitch versus timbre*. Paper presented at the Seventh International Congress on Acoustics, Budapest, 1971.

Portnoy, *The philosopher and music*. New York: The Humanities Press, 1954.

Rameau, J. P. *Traite de l'harmonie reduite a ses principes naturels*. In O. Strunk (Ed.), *Source readings in music history*. New York: Norton, 1950. (Originally published, 1722.)

Rasch, R. A. The perception of simultaneous notes such as in polyphonic music. *Acustica*, 1978, *40*, 1-72.

Restle, F.Theory of serial patterns learning: Structural trees. *Psychological Review*, 1970, *77*, 481-495.

Restle, F., & Brown, E. Organization of serial pattern learning. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol.4). New York: Academic Press, 1970.

Revesz, G. *Zur Grundleguncy der Tonpsychologie*. Leipzig: Feit, 1913.

Risset, J. C., & Mathews, M.V. Analysis of musical instrument tones. *Physics Today*, 1969, *22*, 23-30.

Risset, J. C., & Wessel, D.L. Exploration of timbre by analysis and synthesis. In D. Deutsch (Ed.), *The psychology of music*. New York: Academic Press, 1982.

Rosner, B. S., & Meyer, L.B. Melodic processes and the perception of music, In D, Deutsch (Ed.), *The psychology of music*. New York: Academic Press, 1982.

Ruckmick, C. A. A new classification of tonal qualities. *Psychological Review*, 1929, *36*, 172-180.

Russell, B. *A history of Western philosophy*. New York: Simon and Schuster, 1945.

Saldanha, E. L., & Corso, J. F. Timbre cues for the recognition of musical instruments. *Journal of the Acoustical Society of America*, 1964, *36*, 2021-2026.

Salzer, F. *Structural hearing*. New York: Dover, 1962.

Schenker, H. *Neue musikalische theorien und phantasien: Der freie satz*. Vienna, Austria: Universal Edition, 1956.

Schenker, H. [*Harmony*] (O. Jonas, Ed. & E. M. Borgese, trans.). Cambridge, MA: M.I.T. Press, 1973.

Shoenberg, A. *Harmonielehre*. Leipzig and Vienna: Universal Edition, 1911.

Schoenberg, A. *Style and idea.* London: Williams and Norgate, 1951.

Schouten, J. F. On the perception of sound and speech: Subjective time analysis. *Fourth International Congress on Acoustics, Copenhagen Congress Report II*, 1962, 201-203.

Schroeder, M. R. Acoustics in human communications: Room acoustics, music, and speech. *Journal of the Acoustical Society of America*, 1980, *68*, 22-28.

Shepard, R. N. Circularity in judgments of relative pitch. *Journal of the Acoustical Society of America*, 1964, *36*, 2345-2353.

Simon, H. A. Complexity and the representation of patterned sequences of symbols. *Psychological Review*, 1972, *79*, 369-382.

Simon, H. A., & Kotovsky, K. Human acquisition of concepts for sequential patterns. *Psychological Review*, 1963, *70*, 534-546.

Simon, H. A., & Sumner, R. K. Pattern in music. In B. Kleinmuntz (Ed.), *Formal representation of human judgement*. New York: Wiley, 1968.

Slawson, A. W. Vowel quality and musical timbre as functions of spectrum envelope and fundamental frequency. *Journal of the Acoustical Society of America*, 1968, *43*, 87-101.

Stevens, S. S., & Volkmann, J. The relation of pitch to frequency: A revised scale. *American Journal of Psychology*, 1940, *53*, 329-353.

Strunk, O. (Ed.). *Source readings in music history*. New York: Norton, 1950.

Sutherland, N. S. Object recognition. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. 3). New York: Academic Press, 1973.

Van Noorden, L. P. A. S. *Temporal coherence in the perception of tone sequences.* Unpublished doctoral thesis,

Technische Hogeschool, Eindhoven, Holland, 1975.

Vitz, P. C., & Todd, T. C. A model of learning for simple repeating binary patterns. *Journal of Experimental Psychology*, 1967, *75*, 108-117.

Vitz, P. C., & Todd, T. C. A coded element model of the perceptual processing of sequential stimuli. *Psychological Review*, 1969, *76*, 433-449.

Ward, W. D., & Burns, E. M. Absolute pitch. In D. Deutsch (Ed.), *The psychology of music*. New York: Academic Press, 1982.

Warren, R. M. Auditory temporal discrimination by trained listeners. *Cognitive Psychology,* 1974, *6*, 237-256.

Warren, R. M., Obusek, C. J., Farmer, R. M., & Warren, R. P. Auditory sequence: Confusions of patterns other than speech or music. *Science*, 1969, *164*, 586-587.

Wedin, L., & Goude, G. Dimension analysis of the perception of instrumental timbre. *Scandinavian Journal of Psychology*, 1972, *13*, 228-240.

Werner, H. Uber Mikromelodik und Mikroharmonik. *Zeitschrift fur Psychologie*, 1925, *98*, 74-89.

Wertheimer, M. Untersuchungen sur Lehre von der Gestalt, II. *Psychologische Forschung,* 1923, *4*, 301-350.

Wessel, D. L. Psychoacoustics and music. *Bulletin of the Computer Arts Society*, 1973, *1*, 30-31.

Wessel, D. L. *Low dimensional control of timbre*. IRCAM Report No. 12, 1978, Paris.

Westergaard, P. *An introduction to tonal theory*. New York: Norton, 1975.

White, B. Recognition of distorted melodies. *American Journal of Psychology*, 1960, *73*, 100-107.

Winston, P. H. Learning to identify toy block structures. In R. L. Solso (Ed.), *Contemporary issues in cognitive psychology: The Loyola symposium*. Washington, DC: Winston, 1973.

Yeston, M. (Ed.). *Reading in Schenker analysis and other approaches.* New Haven, CT: Yale University Press, 1977.

Yngve, V. H. A model and an hypothesis for language structure. *Proceedings of the American Philosophical Society*, 1960, *104*, 444-466.

Zarlino, G. *Instituzioni armoniche* (Book 3). In O. Strunk (Ed.), *Source readings in music history*. New York: Norton, 1950. (Originally published, 1558.)