

EXPLORING THE SENSITIVITY TO STRUCTURE IN MUSIC*

Annabel J. Cohen

Music, cornucopia for psychology experiments, abounds with pitch, timbre, intervals, triads, keys, harmony, rhythm, themes, variations, styles, and beauty. Recent technology renders control of these musical elements for study in psychology experiments, and the new mainstream cognitive psychology welcomes music as a route to understanding thought, inference, and imagery. Therefore, music is becoming a more frequent item on the experimental psychologist's menu.

The earlier paucity of musical psychological research furnished a weak base for the present surge of interest. During the Behaviorist epoch the view was prevalent that music was unfit for laboratory consumption! Nevertheless, work on music by those who either preceded Behaviorism, or by those who persisted in spite of it, has assisted the development of a comprehensive theory of the acquisition of sensitivity to musical structure. In particular, two theoretical heritages, seen in the light of each other, can accelerate the attainment of this goal. In the following, the two frameworks are briefly described and research is presented which illustrates their interaction and potential for integration.

Double Heritage

During the past and only century of experimental psychology, two contrasting approaches were applied to problems in

*Research assistance was provided by Mr. Dale Maves, Ms. Barbara Emberly, and Dr. Paul Isaacs, then at the University of Waterloo. The support of the University of Waterloo and the National Research Council is acknowledged. Dr. Lola Cuddy and Dr. Howard Kaplan are thanked for their criticism of an earlier draft of this paper. It was a privilege and honor to make this presentation at the symposium at Queen's University, my Alma Mater.

musical perception. On the one hand, the empiricist tradition emphasized both the role of past experience and the significance of the overtone series. From these first principles, Helmholtz (1885) as a major exponent of the empiricist view, accounted for fundamental musical phenomena such as tonality and chord progression. On the other hand, Gestalt psychologists, Wertheimer (1944), for example, emphasized the spontaneous grouping of elements into perceived wholes. Melodic transposition was taken as a prototypic example because the preservation of musical form relied upon relations among tones rather than upon the tones themselves (see von Ehrenfels 1890).

The phenomena accounted for by the two frameworks co-occur. The transposable melodies of the Gestalt psychologists have Helmholtz's tonality and overtones. Hence, the musical stimulus itself provides a focus for the juxtaposition of the two frameworks. Why such a juxtaposition failed to inspire controversy, research, and creative synthesis is possibly attributable to difficulties inherent in musical-psychological research.

Difficulties in Research in Music Perception

Large structures. The term "music" in everyday usage refers to songs, symphonies, opera, suites, etc. Typical stimuli for psychological tests are much shorter in order to provide time for repeated measurements, controls, and comparisons. Can the smaller unit elucidate the larger work? Musical structure is unique in its exploitation of the same relations at different structural levels. The relations within the smallest part (e.g., the overtone series in a single note, or the repetition of a waveform which gives rise to the perception of pitch) are found in larger units (e.g., the transposition of a theme in sonata form, or the repetition in rhythmic patterns). Perhaps more than for other classes of stimuli, for music the part may reveal aspects of the whole. Nevertheless, certain aspects of the larger structure may never be revealed through the part, and this fact may have discouraged potential researchers.

Understanding the perception of musical microstructure, now feasibly studied with the aid of current technology, may provide useful constraints on the basic assumptions of the theories of the larger work. The results of such studies shed light on the appropriate base level of analysis: waveform, tone, or interval. A decision about first principles of a theory has enormous consequences. In this case, the decision is open to empirical tests of what physical differences are perceptually distinct. Psycholo-

gists who would like to study musical perception might neither like to study the perception of the more amenable musical microstructure, nor appreciate the significance of doing so. Thus the tantalization of long, meaningful musical works which are not easily brought into the laboratory, in contrast to the perhaps less appealing, more controllable waveforms, notes, and intervals may have blocked productivity.

Individual differences. Quite apart from the complexity of music itself, well documented in the centuries of scholarly writing about music, additional difficulties beset the description of the mental representation of musical structure. Individuals differ widely in their musical abilities, depending on age, experience, attention, hearing mechanism, innate intellectual propensities, etc. On separate occasions a given person may perform experimental tasks with varying accuracy or strategy. These difficulties in research can be handled with careful experimental design and statistical analysis. The extraction of general principles from the variation in human responses, and the demonstration of the relation between variables such as age and experience, are most significant contributions to the development of a theory of the acquisition of musical knowledge.

Tacit inference. Another problem in psychological experiments on music is the possible lack of correspondence between subjective and objective musical reality. Music appreciation relies on the mentally generated expectancies of notes soon to be heard, which music theorists and analysts have often tried to describe. See for example, the Preface to Graham George's *Tonality and Musical Structure*, which refers to the perception of tonality that paradoxically is there when it is not there:

Tonalities, on the other hand, do have the capacity to act as boundary materials for the definition of structure. They have it because each of them has implicitly the nature of a thought continuum, and so can be in a sense "there" even when it is not there. So that the return "home" in a tonal structure is not a refashioning of something "like" what was there before (as a return to a thematic procedure is bound to be) but a return to something that has been there — and been "home" even when it was out of sight — all the time (1970: 16-17; italics mine).

Charles Rosen (1981) provides another specific example of perceiving what is not there in reference to Schumann's *Humoreske*, Op. 20. One section comprises three musical staves the middle of

which is denoted "Innere Stimme." This inner melody is implied by the outer bass and treble clefs, the only staves to be played, yet according to Rosen, the *innere stimme* is what Schumann (and he) also heard. Under what circumstances would this melody be heard and recognized by other listeners? In another milieu, jazz musicians claim to hear melodies implied by particular chord patterns and rhythms (see Tirro 1967). Recently, the problem of implication has been explored by Eugene Narmour (1977) in an attempt to accommodate the context established in the minds of listeners while listening to a piece of music. He argues that this fundamental aspect of music appreciation is not accounted for easily by a theory — such as that promoted by Schenker and his followers — which derives musical structure from operations applied to a major triad and a descending scale.

Defining the implications of notes is a task which psychology should be able to assist. After all, implication is very close to a notion of expectation, a concept aligned with empiricist ideas about the role of experience and association in perception. Should not psychological research help to answer questions such as at what point in a musical piece are expectations generated? How soon is a key established? What are the implications of notes having particular scale functions?

Working with short excerpts from the Preludes of the *Well-Tempered Clavier* by J.S. Bach, I explored the expectations of music students. Passages were the first four notes, the first four measures, the first eight measures and the last four measures of the first twelve Preludes. After having heard each excerpt, subjects were asked to sing the first musical scale that came into their minds. In correspondence with musical theoretical notions, the shortest passages, that is, the four-note passages, were sufficient to establish a key. As indicated in Figure 1, the students reliably produced the scale (tonic and mode) of the Prelude when they heard only the opening notes even though some of the notes of the scale were not included and the notes belonged to other scales as well. The results indicated the possibility of quantifying precisely a hierarchy of relations for establishing a sense of key. Related work has been carried out by Helen Lyons and Lola Cuddy (1982) and Carol Krumhansl (1979). Therefore, psychological research has the potential, as yet for the most part unrealized, for providing evidence upon which theories about musical implications can be based.

The interpretation of psychology by the music community. Subtle issues are often the crucial ones for a discipline. The

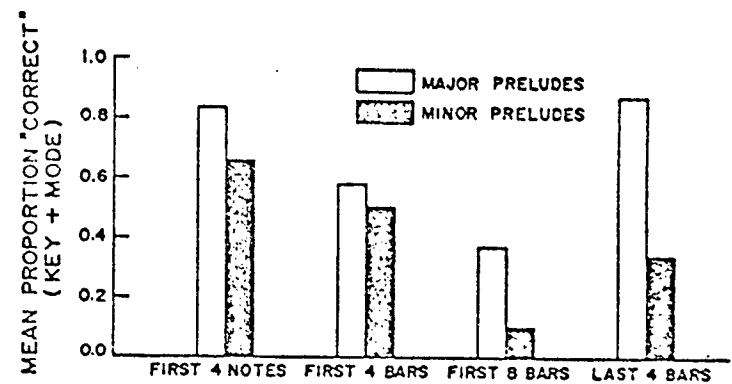


Figure 1

Mean proportion correct vocal production of the tonic plus mode for excerpts of four lengths from six major-mode and six minor-mode Preludes from the Well-Tempered Clavier

subtleties may easily be missed by readers outside the discipline, who, for example, may interpret a presentation of one side of an argument as established fact in the absence of knowledge of the other side. A possible consequence of repeated unchallenged ideas about musical phenomena which dot the past history of psychology is the illusion that more has been solved by psychological research than is indeed the case. Gestalt grouping principles have been accepted as fact rather than treated as a reappearing point of view (see Jackendoff & Lerdahl 1981; Narmour 1977; and Rahn 1981).

In the light of the present discussion, the unresolved and deep controversy between empiricist and Gestalt approaches to perception is an appropriate context for ideas about music from either position (see Herrnstein & Boring 1965: 163 for a discussion of the original philosophical differences). For example, that melodic pattern can be recognized under transposition has been used in support of either a holistic or elementarist approach (compare Gestalt writers with Deutsch 1969). Presumably each approach is incomplete and perhaps incorrect, but where, why, and should there not be some integration of the merits of each?

A Research Problem which Highlights the Intersection of the Theoretical Viewpoints: Helmholtz meets Gestalt

The research project to be reported focuses on the intersection of Gestalt and Helmholtzian approaches to musical perception and questions the role of absolute frequency information in

the recognition of a melody which has been transposed. The Gestalt psychologists claimed that melodies were recognizable when transposed and they provided some qualification that certain melodies might be more easily recognizable than others — those that were the most well-formed and obedient to the laws of organization. (Albert Bregman in the present volume indicates that the characteristics which lead to organization were not described precisely even with respect to visual examples. His paper describes recent progress on the problem.) Gestalt psychology de-emphasized the role of absolute information and overlooked the point that certain degrees of transposition produce greater physical differences from the original sequence than others. Western European music disproportionately exploits transposition to the dominant (see Rosen 1972). For diatonic melodies there is considerable sharing of frequencies from the original to the transposed melody. A tonal melody from the diatonic scale, for example, C D E F G A B, which is transposed from the key of C to the key of G requires the replacement of only one frequency. Because this transposition is the most common in Western European music, the amount of experience with it is confounded with the amount of sharing of pitch information between the original and resulting transposed melody. That is, the transposition most often heard is also the one in which there is the greatest overlap of pitch information.

Helmholtz based his theory of tonality and chord progression on the shared frequencies in chords and in complex tones. As well as tonality, the physical composition of the tones and the degree of sharing of frequencies is important from the Helmholtzian point of view. The role of past musical experience is another variable significant to the empiricist view.

Following from earlier work on exploring sensitivity to rules of tonality in short melodies (see Cuddy, Cohen, & Mewhort 1981), I carried out experiments using two seven-note sequences, one diatonic, the other non-diatonic. As indicated in Figure 2 they were similar in terms of contour and range and began and ended on the same note. The diatonic sequence was characterized by a I-V₇-I progression and the final interval was the leading-tone-to-tonic. For the non-diatonic sequence neither of these two latter features applied.

The paradigm was that of two-alternative forced-choice in which a standard sequence was presented followed by two comparison sequences, one of which is identical to the standard, the

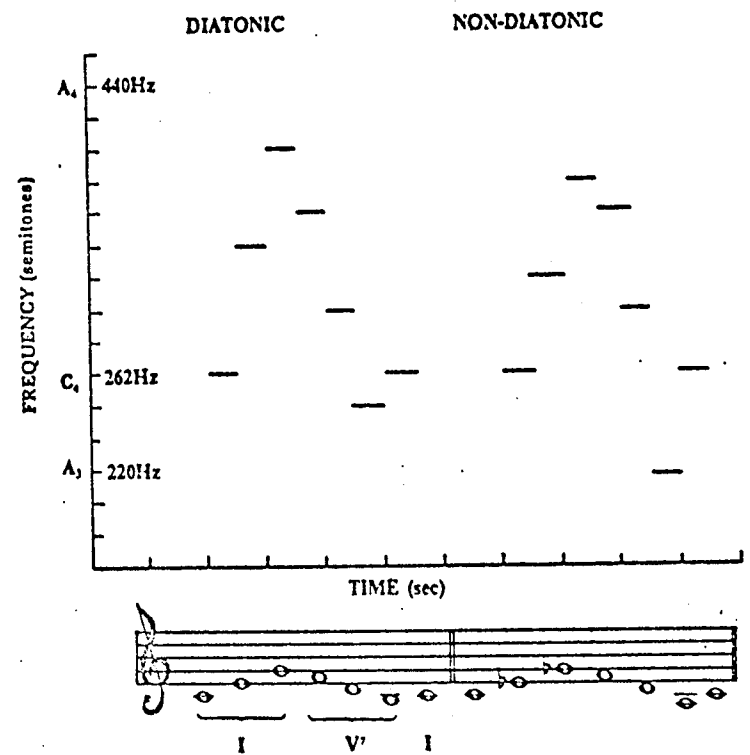


Figure 2

Diatonic (left) and non-diatonic (right) melodies used in the experiment represented in terms of logarithmic frequency (top) and musical notation (bottom)

other of which has been altered (in this case by raising the third or lowering the fifth note by one semitone). Thus, the interval relations were preserved in the correct comparison but were altered in the incorrect comparison. The paradigm has also been used in the study of intervals and triads (see Cuddy & Cohen 1976).

While transposed melodies are identical in terms of intervals, they are similar to varying degrees in terms of the number of musical notes they share with the original melody. In the present example a transposition of the diatonic melody to the tritone would require the replacement of as many notes as would the dominant transposition of the non-diatonic melody, and conversely. It was therefore possible to investigate the interaction of shared frequency, transposition type, and melodic structure. Our earlier work had shown that ability to recognize a transposed sequence depends on its strength of tonality. It also appeared that for our non-diatonic sequences, tritone transposition produced

superior transposed recognition than did dominant transposition, and that for tonal sequences dominant transposition produced higher recognition scores.

In addition, I was interested in the possible role of frequencies shared through the overtone structure. In the present study three different timbres were examined (Figure 3a): (i) a sine tone; (ii) two sine tones related in the ratio of 3:2, the same relation as the perfect fifth interval or of the dominant transposition; and (iii) two sine tones related in the ratio 11:8 approximating the tritone interval of six semitones (the choice of the ratio of 11:8 as opposed to 7:5, the closer approximation to the tritone, was based on the greater ease in producing the same fundamental frequency in corresponding melodies across the timbre conditions, i.e., corresponding notes were related as 8, 11:8, 12:8).

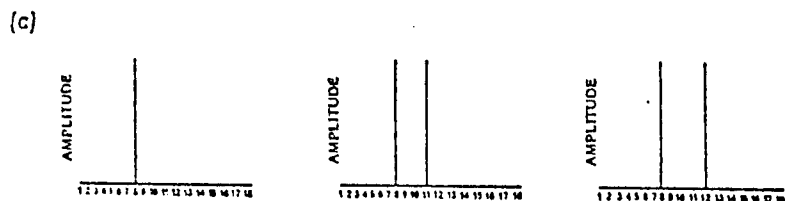
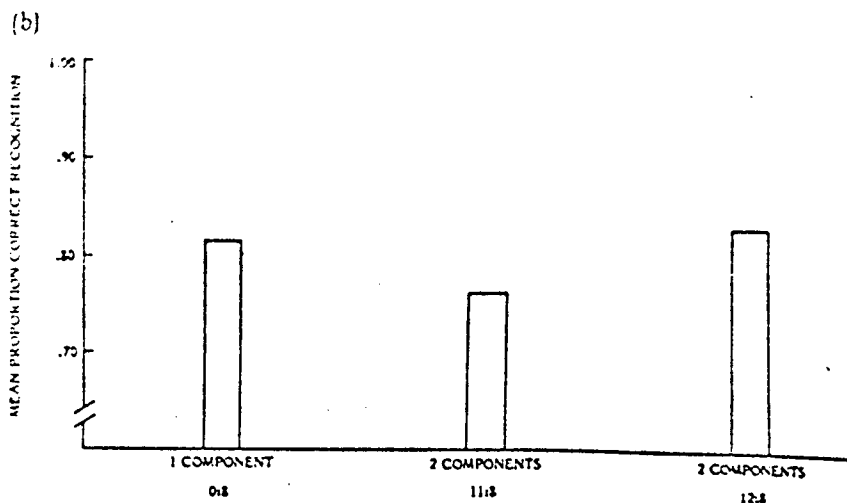


Figure 3

(a) Spectral components of the three timbres:

(b) mean proportion correct recognition as a function of each timbre

The tones were synthesized by the MUSIC-11 program with an "instrument" set for harmonics at 8, 11 and 8, 12 and 8, for the separate timbres. Stimulus tapes were presented to subjects individually from a REVOX A77 tape recorder via headphones in a sound-attenuated room at a loudness level judged comfortable to listeners.

For the timbre with two sine tones related in the ratio 3:2 the upper and lower harmonics of the original and transposed melody coincided under dominant transposition but not under tritone transposition. The timbre constructed of two sine tones related 11:8 produced more overlap with the transposed melody under tritone than under dominant transposition. In addition, greater physical similarity resulted from diatonic sequences formed from the 3:2 timbre and transposed to the dominant, and for the non-diatonic melody formed from the 11:8 timbre and transposed to the tritone than for other combinations of timbres and transpositions. Of course the greatest physical similarity was produced when the comparison sequences were not transposed. Helmholtz's theory predicts, (i) that ease of recognition of the transposed melody would be greatest when there is greatest overall physical similarity between the original and its transposition; (ii) that experienced subjects would be more sensitive in the task both overall and to the timbre and transposition manipulations; and (iii) that the more familiar dominant transposition and tonal (diatonic) melody would be more easily recognized. The strict Gestalt theory would predict that none of these factors should make any difference and that all sequences should be easily recognized under transposition. An alternative interpretation of the theory would predict that the better organized sequences would be more easily recognized when transposed, but what is meant by better organization is not specified.

The response obtained from individual subjects on each trial was entered into an analysis of variance having four within-group factors: first, melodic structure with two levels (diatonic and non-diatonic); second, transposition with three levels (zero, tritone, and dominant); third, timbre with three levels (sine, two components in ratio 11:8, and two components in ratio 12:8), and finally, repetition with four levels corresponding to whether the particular standard sequence was tested for the first, second, third, or fourth time in succession. There was also one between-subjects factor of training with two levels—fifteen highly trained and fifteen less trained subjects. The design is summarized in Table 1.

Number of semitones transposed up

	6	7
	Melody	Melody
	Diatonic	Diatonic
	Non-Diatonic	Non-Diatonic

Timbre	Trial
1 Component	1
	2
	3
	4
2 Components ratio = 11:8 = 1.39	1
	2
	3
	4
2 Components ratio = 12:8 = 1.5	1
	2
	3
	4

Note: 15 subjects with high musical experience and 15 subjects with less training carried out the task.

Table I
Design of the experiment

Musical subjects performed at a higher level than did less experienced subjects with .86 mean proportion correct (PC) for the highly trained listeners vs. .75 PC for the less trained ($F(1,28) = 9.9$, $p < .005$). Both groups found it easier to recognize the diatonic (PC = .88) as opposed to the non-diatonic melody (PC = .74) ($F(1,28) = 42.5$, $p < .0001$). Figure 3b shows a main effect of timbre with the "tritone" timbre producing most difficulty (PC = .76) versus the timbres with components related as the perfect fifth (PC = .83) and the single component (PC = .82) ($F(2,56) = 3.9$; $p < .026$). There was a significant effect of transposition attributable to the greater ease of recognition under no transposition (PC = .85) and more difficulty for the tritone (PC = .79) and dominant (PC = .78) transpositions ($F(2,56) = 8.4$; $p < .001$). The effect of transposition interacted significantly with level of training and melodic structure ($F(2,56) = 3.6$; $p < .03$). As seen in Figure 4, for musically trained subjects the effect of transposition was great for the non-diatonic sequence where the highest score was for no transposition (PC = .83) followed by the tritone transposition (PC = .76) and the dominant transposition (PC = .73). Note that the tritone is slightly superior to the dominant, a finding consistent with the notion of the importance of overlapping pitch information and with earlier work reported by Cuddy, et al. 1981. For the tonal sequences the scores were all high, although the dominant transposition produced the highest score (PC = .97) and the tritone transposition the lowest (PC = .94), again with respect to these two transpositions, consistent with previous work. The less experienced subjects produced the same pattern for the non-diatonic sequence as the more experienced subjects with PC = .74, .68, and .67, for 0, 6, and 7 semitones transposition respectively. For the tonal sequence they showed much greater ease for the 0 transposition (PC = .89) versus the transpositions of 6 and 7 semitones (PC = .76 and .75 respectively). It is notable that for both groups of subjects, the dominant transposition, though in general more familiar, provides no facilitation as compared to the tritone transposition for the non-diatonic sequence. The observation that indeed the tritone transposition was superior here is consistent with the notion of the significance of shared frequency information between transposed and original sequences.

With respect to the effects of varying timbre, facilitation of recognition did not occur under conditions of greatest physical similarity. (Perhaps a closer approximation to 6 semitones for tritone timbre would have enabled this effect.) Nevertheless, an interaction between timbre, transposition, and melodic structure

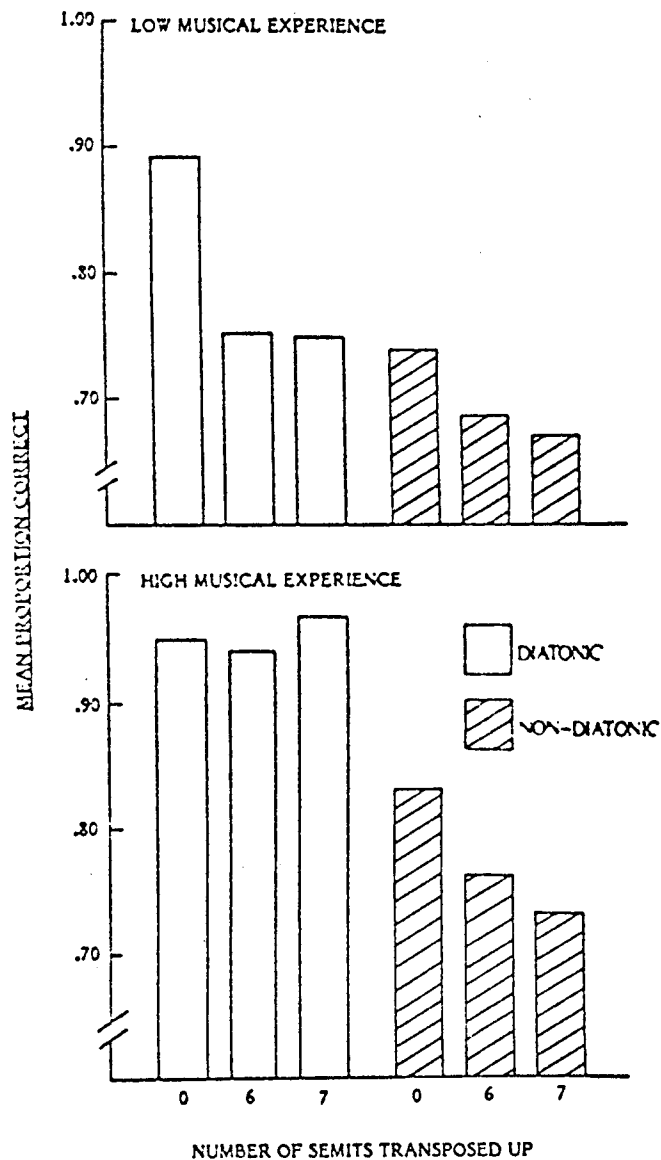


Figure 4

Mean proportion correct recognition as a function of melody, transposition, and musical experience

was evident in the data of the highly trained musicians. Note, with reference to Figure 4 that for the musicians the difference in recognition between diatonic and non-diatonic sequences was greatest for dominant transpositions. While this difference occurred for all three timbres, it was greatest for the timbre with components related 3:2. The data is summarized in Table II, which reports mean difference in proportion correct between diatonic and non-diatonic sequences for all combinations of timbre levels and transposition levels. Table II shows that the greatest difference in recognition of the transposed diatonic and non-diatonic melodies (PC = .25) occurs under the timbre of 3:2 under dominant transposition. Under tritone transposition the difference is reduced (PC = .16). The timbre of 11:8 and the single component timbre give rise to a smaller range of differences (PC = .23 for dominant transposition; PC = .18 for tritone transposition). It is interesting to note that the effects of a simple and complex timbre both given rise to less effect of the musical structure (melodic and transpositional) than the timbre of 3:2 which in some respects falls midway between the other two timbres on the continuum of complexity.

Timbre	Transposition (semitones)		
	0	6	7
1 component	.10	.18	.23
2 components ratio=1.39	.13	.18	.23
2 components ratio=1.5	.12	.16	.25*

Mean proportion correct recognition for the non-diatonic sequence is subtracted from the mean proportion correct recognition for the diatonic sequence.

*Greatest difference.

Table II

Mean difference in difficulty of recognition of diatonic and non-diatonic sequences for each of three timbres under three transpositions for 15 highly experienced listeners

Therefore the abstraction of melodic structure may depend in part upon timbre. In the present study any effect is subtle. Further studies using the present paradigm and other paradigms, for example, multidimensional scaling, will enable the precise specification of the information mentally represented in the perception and memory of melodies. Because melodies contain the same units of analysis as larger works (e.g., the progression, V₇-I), greater understanding of so-called real music may follow.

The present study employed melodic stimuli which contained features which had been discussed by Gestalt and empiricist psychologists, different features in each case. Progress in understanding musical perception may advance by examining the same phenomena from different perspectives rather than by aligning different phenomena with different theoretical perspectives as was the past custom. In this way a better composite framework can develop. For example, Helmholtzian notions may aid the clarification of the Gestalt notion of well-formedness, while the Gestalt emphasis on transposition brings to light the limitations in a theory which emphasizes frequency components and experience.

All of the variables manipulated in the present study, melody structure, transposition, timbre, and training had significant effects on the recognition of transposed melodies. A future study in which the "tritone" timbre contains the same relation as the tritone interval of transposition (1.41 rather than 1.39) may produce the expected facilitation of recognition of non-diatonic sequences in accordance with the Helmholtzian view. If not, then the overtone structure is dominated by structures at higher levels rather than the reverse. Further investigation of the phenomenon, though extremely subtle, is warranted. After all, subtlety is at the core of the appeal of music.

Conclusion

An appreciation of the heritages of psychological research helps guide the application of psychology to music theory. The attempt to resolve theoretical conflicts in psychology using musical materials would facilitate the communication both of psychology to music theorists and of the significance of music theory to psychologists. This endeavor may be nurtured by drawing parallels between the controversies within psychology (Gestalt/empiricist) and within music (Schenker/Narmour).

REFERENCES

- CUDDY, L.L. and COHEN, A.J.
1976: "Recognition of Transposed Melodic Sequences." *Quarterly Journal of Experimental Psychology*, XXVIII/2, 255-70.
- CUDDY, L.L., COHEN, A.J., and MEWHORT, D.J.K.
1981: "Perception of Structure in Short Melodic Sequences," *Journal of Experimental Psychology: Human Perception and Performance*, VII/4, 869-83.
- DEUTSCH, D.
1969: "Music Recognition," *Psychological Review*, LXXVI/3, 300-07.
- EHRENFELS, C. von.
1890: "Über Gestaltqualitäten," *Vierteljahrschrift für wissenschaftliche Philosophie*, XIV, 249-92; see also "On Gestalt-Qualities," translated by Mildred Focht, *Psychological Review*, XLIV/6 (1937), 521-24.
- GEORGE, G.
1970: *Tonality and Musical Structure*. London: Faber & Faber.
- HELMHOLTZ, H.
1954: *On the Sensations of Tone as a Physiological Basis for the Theory of Music*. 4th edition (1877/1885). Translated by A.J. Ellis. New York: Dover.
- HERRNSTEIN, R.J. and BORING, E.G., eds.
1965: *A Source Book in the History of Experimental Psychology*. Cambridge, Mass.: Harvard University Press.
- JACKENDOFF, R. and LERDAHL, F.
1981: "Generative Music Theory and its Relation to Psychology," *Journal of Music Theory*, XXV/1, 45-90.
- KRUMHANSL, C.L.
1979: "The Psychological Representation of Musical Pitch in a Tonal Context," *Cognitive Psychology*, XI/2, 346-74.
- LYONS, H.I. and CUDDY, L.L.
1982: "Musical Modulation and the Perceptual Structure of Tone Sequences," paper presented at the 103rd Meeting of the Acoustical Society of America, Chicago, April 1982.
- NARMOUR, E.
1977: *Beyond Schenkerism: The Need for Alternatives in Music Analysis*. Chicago: University of Chicago Press.
- RAHN, J.
1982: "Where is the Melody?" *In Theory Only*, VII/6, 3-19.
- ROSEN, C.
1972: *The Classical Style: Haydn, Mozart, Beethoven*. New York: W.W. Norton.
1981: "The Fragment as an Element in Nineteenth-Century Music." C.A.P.A.C./MacMillan Lectures, Toronto, 1 ne.

TIRRO, F.

1967: "The Silent Theme Tradition in Jazz." *The Musical Quarterly*,
LIII/3, 313-34.

WERTHEIMER, M.

1944: "Gestalt Theory." translated by N. Nairn-Allison, *Social
Research*, XI/1, 81-99; excerpted in W.D. Ellis, ed., *A Source
Book of Gestalt Psychology*. New York: Harcourt, Brace, 1939.
1-11.