1995

The Perception Of Rhythm And Tempo Modulation In Music

Matthew Stratton Royal

Recommended Citation

Royal, Matthew Stratton, "The Perception Of Rhythm And Tempo Modulation In Music" (1995). *Digitized Theses*. 2535.
https://ir.lib.uwo.ca/digitizedtheses/2535

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

THE PERCEPTION OF RHYTHM AND TEMPO MODULATION IN MUSIC

by

Matthew S. Royal

Faculty of Music

Submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy

Faculty of Graduate Studies
The University of Western Ontario
London, Ontario
April 1995

National Library of Canada
Acquisitions and Bibliographic Services Branch
395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada
Direction des acquisitions et des services bibliographiques
395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

ISBN 0-612-03486-0

Name MATTHEW ROYAL

Dissertation Abstracts International is arranged by broad, general subject categories. Please select the one subject which most nearly describes the content of your dissertation. Enter the corresponding four-digit code in the spaces provided.

MUSIC
SUBJECT TERM

0413
SUBJECT CODE

U·M·I

Subject Categories

THE HUMANITIES AND SOCIAL SCIENCES

COMMUNICATIONS AND THE ARTS

Architecture	0729
Art History	0377
Cinema	0900
Dance	0378
Fine Arts	0357
Information Science	0723
Journalism	0391
Library Science	0399
Mass Communications	0708
Music	0413
Speech Communication	0459
Theater	0465

EDUCATION

General	0515
Administration	0514
Adult and Continuing	0516
Agricultural	0517
Art	0273
Bilingual and Multicultural	0282
Business	0688
Community College	0275
Curriculum and Instruction	0727
Early Childhood	0518
Elementary	0524
Finance	0277
Guidance and Counseling	0519
Health	0680
Higher	0745
History of	0520
Home Economics	0278
Industrial	0521
Language and Literature	0279
Mathematics	0280
Music	0522
Philosophy of	0998
Physical	0523
Psychology	0525
Reading	0535
Religious	0527
Sciences	0714
Secondary	0533
Social Sciences	0534
Sociology of	0340
Special	0529
Teacher Training	0530
Technology	0710
Tests and Measurements	0288
Vocational	0747

LANGUAGE, LITERATURE AND LINGUISTICS

Language	
General	0679
Ancient	0289
Linguistics	0290
Modern	0291
Literature	
General	0401
Classical	0294
Comparative	0295
Medieval	0297
Modern	0298
African	0316
American	0591
Asian	0305
Canadian (English)	0352
Canadian (French)	0355
English	0593
Germanic	0311
Latin American	0312
Middle Eastern	0315
Romance	0313
Slavic and East European	0314

PHILOSOPHY, RELIGION AND THEOLOGY

Philosophy	0422
Religion	
General	0318
Biblical Studies	0321
Clergy	0319
History of	0320
Philosophy of	0322
Theology	0469

SOCIAL SCIENCES

American Studies	0323
Anthropology	
Archaeology	0324
Cultural	0326
Physical	0327
Business Administration	
General	0310
Accounting	0272
Banking	0770
Management	0454
Marketing	0338
Canadian Studies	0385
Economics	
General	0501
Agricultural	0503
Commerce-Business	0505
Finance	0508
History	0509
Labor	0510
Theory	0511
Folklore	0358
Geography	0366
Gerontology	0351
History	
General	0578
Ancient	0579
Medieval	0581
Modern	0582
Black	0328
African	0331
Asia, Australia and Oceania	0332
Canadian	0334
European	0335
Latin American	0336
Middle Eastern	0333
United States	0337
History of Science	0585
Law	0398
Political Science	
General	0615
International Law and Relations	0616
Public Administration	0617
Recreation	0814
Social Work	0452
Sociology	
General	0626
Criminology and Penology	0627
Demography	0938
Ethnic and Racial Studies	0631
Individual and Family Studies	0628
Industrial and Labor Relations	0629
Public and Social Welfare	0630
Social Structure and Development	0700
Theory and Methods	0344
Transportation	0709
Urban and Regional Planning	0999
Women's Studies	0453

THE SCIENCES AND ENGINEERING

BIOLOGICAL SCIENCES

Agriculture	
General	0473
Agronomy	0285
Animal Culture and Nutrition	0475
Animal Pathology	0476
Food Science and Technology	0359
Forestry and Wildlife	0478
Plant Culture	0479
Plant Pathology	0480
Plant Physiology	0817
Range Management	0777
Wood Technology	0746
Biology	
General	0306
Anatomy	0287
Biostatistics	0308
Botany	0309
Cell	0379
Ecology	0329
Entomology	0353
Genetics	0369
Limnology	0793
Microbiology	0410
Molecular	0307
Neuroscience	0317
Oceanography	0416
Physiology	0433
Radiation	0821
Veterinary Science	0778
Zoology	0472
Biophysics	
General	0786
Medical	0760

EARTH SCIENCES

Biogeochemistry	0425
Geochemistry	0996
Geodesy	0370
Geology	0372
Geophysics	0373
Hydrology	0388
Mineralogy	0411
Paleobotany	0345
Paleoecology	0426
Paleontology	0418
Paleozoology	0985
Palynology	0427
Physical Geography	0368
Physical Oceanography	0415

HEALTH AND ENVIRONMENTAL SCIENCES

Environmental Sciences	0768
Health Sciences	
General	0566
Audiology	0300
Chemotherapy	0992
Dentistry	0567
Education	0350
Hospital Management	0769
Human Development	0758
Immunology	0982
Medicine and Surgery	0564
Mental Health	0347
Nursing	0569
Nutrition	0570
Obstetrics and Gynecology	0380
Occupational Health and Therapy	0354
Ophthalmology	0381
Pathology	0571
Pharmacology	0419
Pharmacy	0572
Physical Therapy	0382
Public Health	0573
Radiology	0574
Recreation	0575
Speech Pathology	0460
Toxicology	0383
Home Economics	0386

PHYSICAL SCIENCES

Pure Sciences

Chemistry	
General	0485
Agricultural	0749
Analytical	0486
Biochemistry	0487
Inorganic	0488
Nuclear	0738
Organic	0490
Pharmaceutical	0491
Physical	0494
Polymer	0495
Radiation	0754
Mathematics	0405
Physics	
General	0605
Acoustics	0986
Astronomy and Astrophysics	0606
Atmospheric Science	0608
Atomic	0748
Electronics and Electricity	0607
Elementary Particles and High Energy	0798
Fluid and Plasma	0759
Molecular	0609
Nuclear	0610
Optics	0752
Radiation	0756
Solid State	0611
Statistics	0463

Applied Sciences

Applied Mechanics	0346
Computer Science	0984
Engineering	
General	0537
Aerospace	0538
Agricultural	0539
Automotive	0540
Biomedical	0541
Chemical	0542
Civil	0543
Electronics and Electrical	0544
Heat and Thermodynamics	0348
Hydraulic	0545
Industrial	0546
Marine	0547
Materials Science	0794
Mechanical	0548
Metallurgy	0743
Mining	0551
Nuclear	0552
Packaging	0549
Petroleum	0765
Sanitary and Municipal	0554
System Science	0790
Geotechnology	0428
Operations Research	0796
Plastics Technology	0795
Textile Technology	0994

PSYCHOLOGY

General	0621
Behavioral	0384
Clinical	0622
Developmental	0620
Experimental	0623
Industrial	0624
Personality	0625
Physiological	0989
Psychobiology	0349
Psychometrics	0632
Social	0451

ABSTRACT

Research is presented on the perception of rhythm, specifically, the detection of changing or "modulating" inter-onset intervals in simple musical stimuli. The types of changes of rhythmic patterns examined are "rhythm modulation" and "tempo modulation." These two terms are akin to the musician's concepts of agogics, rubato, accelerando and ritardando, all being common expressive devices in musical performance. Rhythm modulation occurs when an initially even or isochronous rhythm becomes increasingly more uneven. Tempo modulation occurs when the beat rate of a rhythm accelerates or decelerates.

Two approaches are adopted to elucidate how such modulating patterns might be perceived, a theoretical one and an experimental one. Following a review of the pertinent literature, a theoretical model of time-interval perception in music is proposed that attempts to synthesize the findings of previous experimentation. The main thrust of the model is that rhythm perception is mediated by two complementary processes: (1) a so-called OSCILLATOR BANK that entrains to stimulus time-intervals on a note-to-note basis, and (2) a SHORT AUDITORY STORE that is responsible for integrating temporally separated events.

The model generates the hypothesis that rhythm modulation will be detected in the OSCILLATOR BANK, whereas tempo modulation will be detected in the SHORT AUDITORY STORE. This hypothesis is tested in three perceptual experiments. To compare the difficulty of detecting rhythm and tempo modulation under various conditions, certain variables are manipulated: the direction of modulation (whether a change onset occurs earlier or later than expected), the initial beat rate, the metrical location of modulation, and the presence or absence of beat subdivision. To measure perceptual difficulty, a type of reaction-time dependent variable and a modulation-type-identification dependent variable are used.

The following results are observed: (a) the direction of modulation is significant only for tempo modulation, (b) rhythm and tempo modulation exhibit contrasting trends across the musical initial-beat-rate range, (c) metrical location does not affect detection and (d) detection is easier with beat subdivision. These results are generally consistent with the hypothesis that rhythm and tempo modulation detection are mediated by contrasting perceptual processes.

To my parents, for years of love and support

ACKNOWLEDGMENTS

I should like to acknowledge with gratitude the following people who, in various ways, have helped this dissertation come to fruition:

My advisor, Professor Harold Fiske, whose sagacity, insight, good humour and patience have been constant throughout my time under his tutelage. I am indebted to Professor Fiske for two intellectually provoking graduate courses in music psychology and experimental design and statistics, and for numerous conversations on the finer points of putting together a research study. From Professor Fiske I have learned the value of gaining solid philosophical and theoretical ground before embarking on empirical work, and without his counsel the research reported here would have taken a rockier road than it has.

My advisory committee, Professors Lloyd Dawe and David Huron, whose pertinent critiques of earlier drafts of this dissertation have been invaluable. To Professor Dawe I extend my thanks for all the help he has freely given concerning certain statistical matters, and for highlighting areas in the psychophysical literature where my own knowledge was wanting. To Professor Huron I express my appreciation for the many stimulating discussions on empirical approaches to music theory and on music psychology in general. The precision with which both scholars were able to root out any woolly thinking or terminological imprecision on my part was exemplary.

Dr. Ian Cross, the supervisor of my Master's dissertation, for igniting the initial spark of enthusiasm for music psychology. Without his openness and engagement in the subject, I should never have started on research in music perception in the first place.

The staff, faculty and my fellow students in the Faculty of Music at the University of Western Ontario for providing an atmosphere that was friendly and welcoming, yet conducive to intellectual endeavour. I am grateful for the ease with which I have had access to an excellent library and to the digital sound laboratory where the experimental stimuli were generated and recorded.

Ms. Karen Lee, formerly of the Department of Statistics and Actuarial Science at the University of Western Ontario, for help and advice with Survival Analysis.

The Faculty of Graduate Studies at the University of Western Ontario, who generously awarded me a Graduate Research Scholarship for the academic year 1993-1994.

Astrid Heyer for constant love, companionship and encouragement when the demands of research made me both buoyant and despondent.

I am much obliged to all of the above who have made my study easier, richer and more enjoyable than it would otherwise have been. Their influences on this work have only been beneficial. Any infelicities that remain are solely the responsibility of the author.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

Terms and Definitions

In their seminal work on rhythm in music, Grosvenor Cooper and Leonard Meyer make the following claim about the centrality of rhythm to musical understanding:

> To study rhythm is to study all of music. Rhythm both organizes, and is itself organized by, all the elements which create and shape musical processes. (Cooper and Meyer 1960, 1)

Similarly, Wallace Berry asserts that,

> All element-processes are rhythmic. In an important sense, the study of rhythm is thus the study of all musical elements, the actions of those elements producing the effects of pace, pattern, and grouping which constitute rhythm. (Berry 1987, 301)

For both Cooper and Meyer and Berry, rhythm is a general term that encompasses such concepts as tempo, grouping, meter, pulse, patterning, activity and so on. Whilst these effects might arise from the ordering and interaction of the diverse elements of musical sound, they all share a relatedness with the durational or temporal organization of musical events. Therefore, a necessary, if not sufficient, condition for any examination of musical rhythm, is that it concern itself with temporal spans between events and groups of events.

Of course, a vast range of physical temporal spans are routinely used in some manner in music-making: a sine wave of 20,000 Hz (roughly the ultrasonic threshold for a healthy young adolescent) has a period of

50 microseconds, whilst Wagner's *Ring* cycle, if performed continuously, might last as long as 20 hours. However, the range of temporal spans that are normally regarded as "rhythmic" is much narrower. The lower bound of this range may be set at around 50 milliseconds (i.e., 1/20 of a second), which is approximately the minimum interval needed for the perception of event succession (Handel 1989, 385). Events occurring faster than 20 times per second fuse together as a single percept, a tone.

The upper bound of the rhythmic temporal-span range is more debatable. Cooper and Meyer, for example, in emphasizing the hierarchical nature of rhythmic organization in music, state,

> Most of the music with which we shall be concerned is architectonic in its organization. That is, just as letters are combined into words, words into sentences, sentences into paragraphs, and so on, so in music individual tones become grouped into motives, motives into phrases, phrases into periods, etc. (Cooper and Meyer 1960, 2)

Cooper and Meyer take this hierarchical view of rhythmic structure to its extreme: whole movements are analyzed as instantiating a single rhythmic grouping (e.g., a trochee or an anapest). Perusal of these authors' analysis of the first movement of Beethoven's Eighth Symphony serves to illustrate their methodology (Cooper and Meyer 1960, 188-203).

In contrast to Cooper and Meyer, others have argued that for a stimulus to be truly rhythmic, its temporal relations must be apprehensible within a single span of attention, within what is termed the *psychological present* (Clarke 1987b; Handel 1989, 385). The psychological present has an assumed span of no more than about 8 seconds and is typically around 2-3 seconds (Michon 1978). The viewpoint adopted in this dissertation is that, while temporal proportions larger than 8 seconds or so can certainly be of importance to the temporal structure of a piece, their reality is *conceptual* rather than *perceptual*. Of course, it is entirely possible that concepts can influence rhythm perception: for example, having the abstract knowledge that a piece exhibits rhythmic symmetry over spans of hundreds of bars may affect the way one chooses to listen that piece. However, in the interests of circumscribing the scope of this dissertation to manageable proportions, and given the perceptual focus of this study, the definition of rhythm employed here will limit itself to immediately perceptible temporal relationships.

Having limited the range within which a temporal span may be said to have a rhythmic quality, one is now ready for a more refined and workable definition of rhythm. Despite the refinement, however, the definition of rhythm adopted in this dissertation is still generic and wide-ranging:

- *Rhythm* is defined as any arrangement of temporal intervals between events in an auditory pattern that may be apprehended in one span of attention by the human listener.

This general definition admits a wide variety of arrangements of temporal intervals from a simple regular pulse, to complex and ambiguous groupings such as polyrhythms or syncopated rhythms. It is worth noting that rhythms in this generic sense do not have to be notatable in musical notation, nor do they necessarily have to incorporate any strict repetition or regularity. An auditory stimulus may be said to be rhythmic insofar as its temporal spans are organized by a perceiver over time.

The above definition outlines in the broadest terms the main problem area of this dissertation, the organization and grouping of immediately apprehensible temporal spans in music. Two particular facets of temporal spans in music will be scrutinized in detail: rhythm modulation and tempo modulation. Rhythm modulation and tempo modulation, along with a number of other related terms are defined in the ensuing paragraphs.

Having established the range of temporal intervals that constitutes a rhythm, one might next ask "intervals between what?" As the above quotations from Cooper and Meyer and Berry intimate, temporal intervals in music may occur between a multitude of different auditory events: pitched notes, unpitched sounds, dynamic stresses, chords, sounds of a particular timbre, prominent pitches within a melody, repetitions of a short melodic figure and so on. For the purposes of this dissertation, all event types may be encompassed by one or both of the following terms:

- An *Event Onset* marks the point in time where a single auditory event (a sound) starts.

Thus, an event onset is the physical rather than perceptual start of a musical event. Since, in the present study, all the sounds considered have rapid physical attack rates (< 100 milliseconds), the difference between physical and perceptual onsets is negligible, and the two may be considered as synchronous (see, however, Vos and Rasch 1981).

- *Accents* are treatments of musical events that make those events perceptually more salient or noticeable than surrounding events.

Normally, the increased salience is achieved by virtue of an event being positively differentiated from its neighbours along one or more of the dimensions of sound. Thus, for example, an accent may arise through increased dynamic level,[1] increased duration, increased timbral richness, melodic prominence, and so on, imparted to a particular event relative to the surrounding events.

Event onsets and accents give rise to two corresponding types of temporal intervals, both of which are illustrated in Figure 1.1:

- *Inter-Onset Interval* (abbreviated as IOI) refers to the temporal interval that elapses between the onsets of two successive auditory events.

The two events bounding an IOI are not necessarily similar types of events. By convention, IOIs are usually measured in milliseconds (hereafter abbreviated as "ms").

- *Inter-Accent Interval* (abbreviated as IAI) refers to the time interval that elapses between two successive accents of the same type.

For example, an IAI could be the time interval between two successive dynamic accents, but not the time interval, say, between a dynamic accent and a melodic accent. Accents of one particular type do not usually occur at successive events but normally have at least one unaccented event intervening. For this reason, an IAI will often encompass several IOIs within its time-span. For the sake of consistency, however, IAIs are also denoted in milliseconds in this dissertation.

---

[1] The term *dynamic* is here used in its musicological sense of "pertaining to loudness or intensity," rather than its psychological sense of "changing" or "in flux."

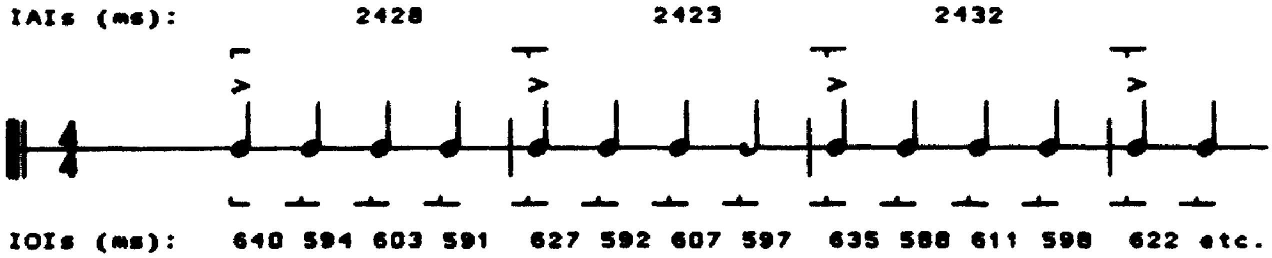


Figure 1.1. Rhythm illustrating *Inter-Onset Intervals* (IOIs) and *Inter-Accent Intervals* (IAIs), both in ms. The note-heads can represent any musical event (e.g., note, chord, unpitched note). The accents, in this case, represent dynamic stresses. The departures from the strict quarter-note rhythm are the type of deviations that might be produced by a human performer.

Despite the ubiquity of the aforementioned accent types in music, and therefore the prevalence of inter-accent durations, the convention employed in this dissertation for identifying a particular rhythm, or of distinguishing between two different rhythms, will be by referring to IOIs. In fact, in much of the literature reviewed in Chapter 2, rhythms are most often classified according to the proportions of successive IOIs. One crude but useful distinction drawn in this study is illustrated in Figure 1.2 and is defined as follows:

- *Isochronous Rhythms* present a constant, unchanging IOI.

The ratio of successive IOIs in an isochronous rhythm is always 1:1. Examples of isochronous rhythms are given in Figures 1.2a and 1.2b.

- *Uneven Rhythms* are rhythms that contain at least two contrasting IOI values (i.e., they are *not* isochronous).

This study makes use of a specific subset of uneven rhythms, namely *repetitive uneven rhythms*, and these are exemplified in Figures 1.2c through 1.2e. The *degree of unevenness* of a rhythm is the extent to which contrasting IOIs depart from a ratio of 1:1, so, for instance, Figure 1.2d is more uneven than Figure 1.2c.

A rhythm may change its degree of unevenness, or *modulate*. For example an isochronous rhythm may become anisochronous due to lengthening or shortening of certain tones. By the same token, an initially uneven rhythm may become less uneven (move towards isochrony) or become still more uneven. An example of what will be termed a *modulating rhythm* is given in Figure 1.2f. The above description of changing unevenness leads to a further definition:

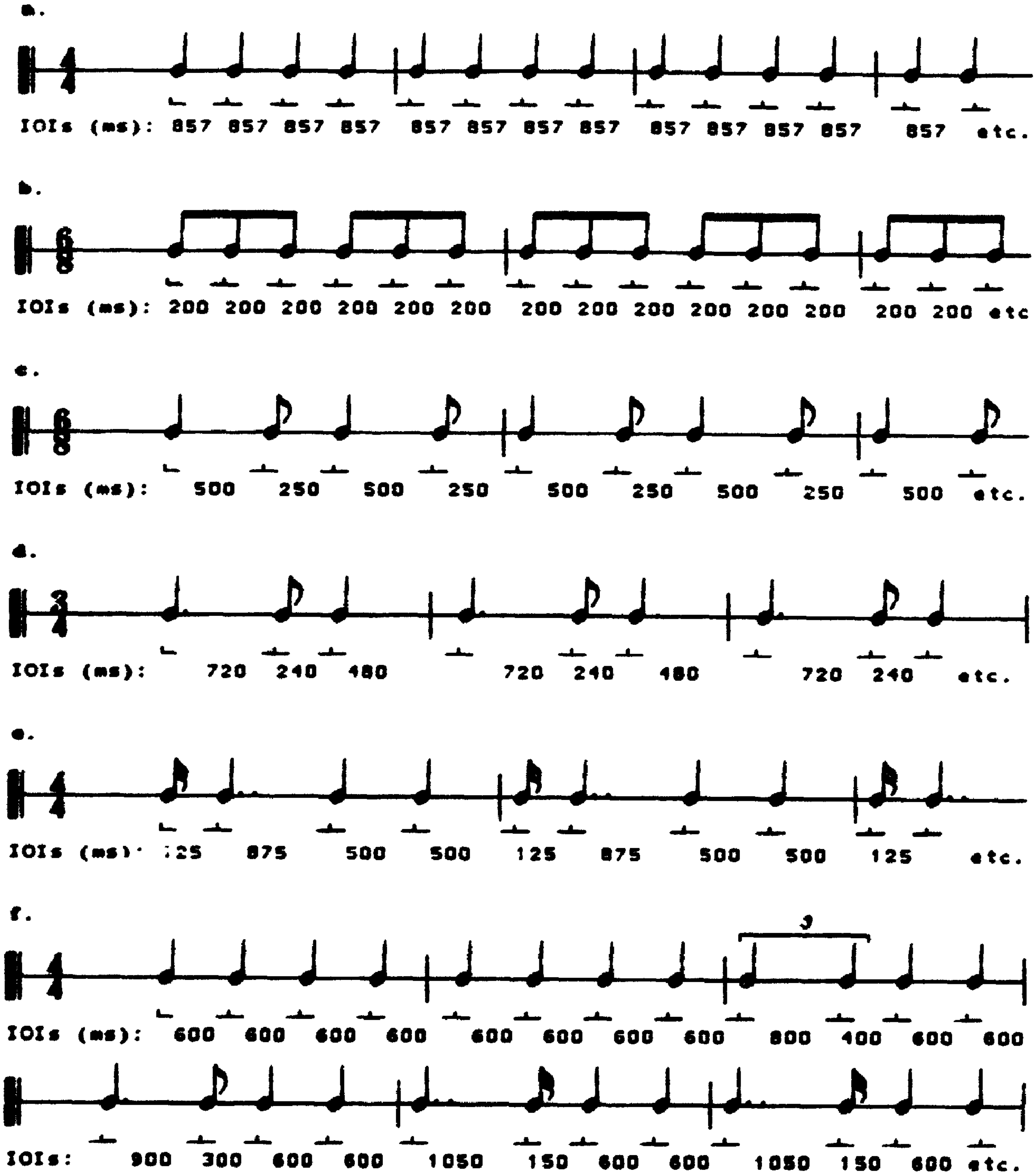


Figure 1.2. Isochronous and repetitive uneven rhythms. Figures a and b show different isochronous rhythms. Figures c, d and e show repetitive uneven rhythms with various degrees of unevenness. Figure f shows a rhythm that modulates from isochrony to repetitive unevenness over the course of three bars.

- *Rhythm Modulation* is defined as a change in the degree of unevenness of a rhythm.

Naturally, the types of rhythms that can occur in music are, in practical terms, limitless and in most cases are more complex than those shown in Figure 1.2. However, repetitive and modulating rhythms are useful for the purposes of perceptual experimentation since they allow one to measure the point at which an increasingly uneven rhythm is no longer categorized by the listener as isochronous.

In order to facilitate the perception of the more complex types of rhythms found in real music, many listeners are able psychologically to abstract a regular isochronous pulse or beat that provides a temporal yardstick against which durations in the musical stimulus are measured. Anyone who has spontaneously tapped his/her foot to a popular tune on the radio will have an intuitive sense of the concept of pulse or beat (here the two terms are taken to be interchangeable). Given this observation, two more terms require formal definition:

- *Beat* is the isochronous repetition of a single, focal time interval, which is perceptually inferred by the listener, and around which time intervals in the musical stimulus are perceptually organized.
- *Tempo* is here defined as the rate at which successive beats occur.

By convention in music, tempo is usually denoted in beats per minute (hereafter abbreviated as "bpm").

Since, by definition, a beat is an isochronous pulse abstracted from a rhythm, a constant beat rate or tempo may be expressed not only in beats per minute, but also as a single *inter-beat interval* (hereafter abbreviated as "IBI") value in milliseconds. Tempo values in beats per minute and IBI values in milliseconds are therefore related by the following formulae:

(1.1) $$\text{IBI} = 60000/\text{Tempo}$$ or

(1.2) $$\text{Tempo} = 60000/\text{IBI}.$$

Thus, for example, a tempo of 60 bpm will have beats 60000/60 or 1000 ms apart. Conversely, beats separated by, say, 750 ms will have a tempo of 60000/750 or 80 bpm.

In a perfectly isochronous rhythm with no subdivision of the beat the physical inter-onset interval (IOI) is very likely to be equal to the psychological inter-beat interval (IBI). In contrast, when a rhythm is complex, changing, or when it presents various subdivisions of the beat, the IOI value will be in a state of flux and will not necessarily be equal to the IBI. In most musical situations, the time intervals between successive events will be highly variable depending on the smallest IOIs present in a rhythm at any given time. To reflect the importance of the often varying rate of successive musical events, a term is adopted from Berry (1987):

- *Rhythmic activity* is the quality of rhythmic motion or "busy-ness," determined largely by the rate of the fastest-occurring events at any given moment.

A perfectly isochronous beat and therefore a perfectly steady tempo are only feasible with computerized performances of music. With all human performers the beat rate will undergo change during the performance, both as a result of error, and deliberately as a feature of expressive playing. For this reason, change of tempo, however subtle and however localized, is an inevitable part of human musical performance. Given the ubiquity of beat-rate fluctuations, a further term requires definition:

- *Tempo modulation* is defined as a change in beat rate.

Tempo modulation may be manifested as either a *decrease* in tempo (deceleration, or lengthening of IBIs), an *increase* in tempo (acceleration or shortening of IBIs), or as sequential combinations of the two. Tempo decrease or deceleration is exemplified in Figure 1.3a; tempo increase or acceleration is exemplified in Figure 1.3b. In addition, the *rate* of tempo modulation may vary. Changes in beat rate can occur instantaneously (when, for example, the beat rate triples between two adjacent notes), or change may evolve more gradually over a number of beats. In the study described here, tempo will modulate in a gradual and linear manner (i.e., by a fixed percentage of the starting IOI over a fixed number of beats) in one direction only.

In the cases of both rhythm and tempo modulation, the temporal change cannot be apparent to the listener until at least the first altered IOI has sounded. Therefore the first sign of rhythm or tempo modulation is the onset that marks the conclusion of the first altered IOI:

- The *First Anomalous Onset* (abbreviated as FAO) is defined as the first onset that occurs earlier or later than would be expected from a previously established rhythmic pattern.

Having introduced the concept of beat, the next logical progression is to *meter*. Meter, like other well-worn terms in musical discourse, may be variously defined depending on the style of music being studied and the theoretical standpoint of the writer. Berry, for

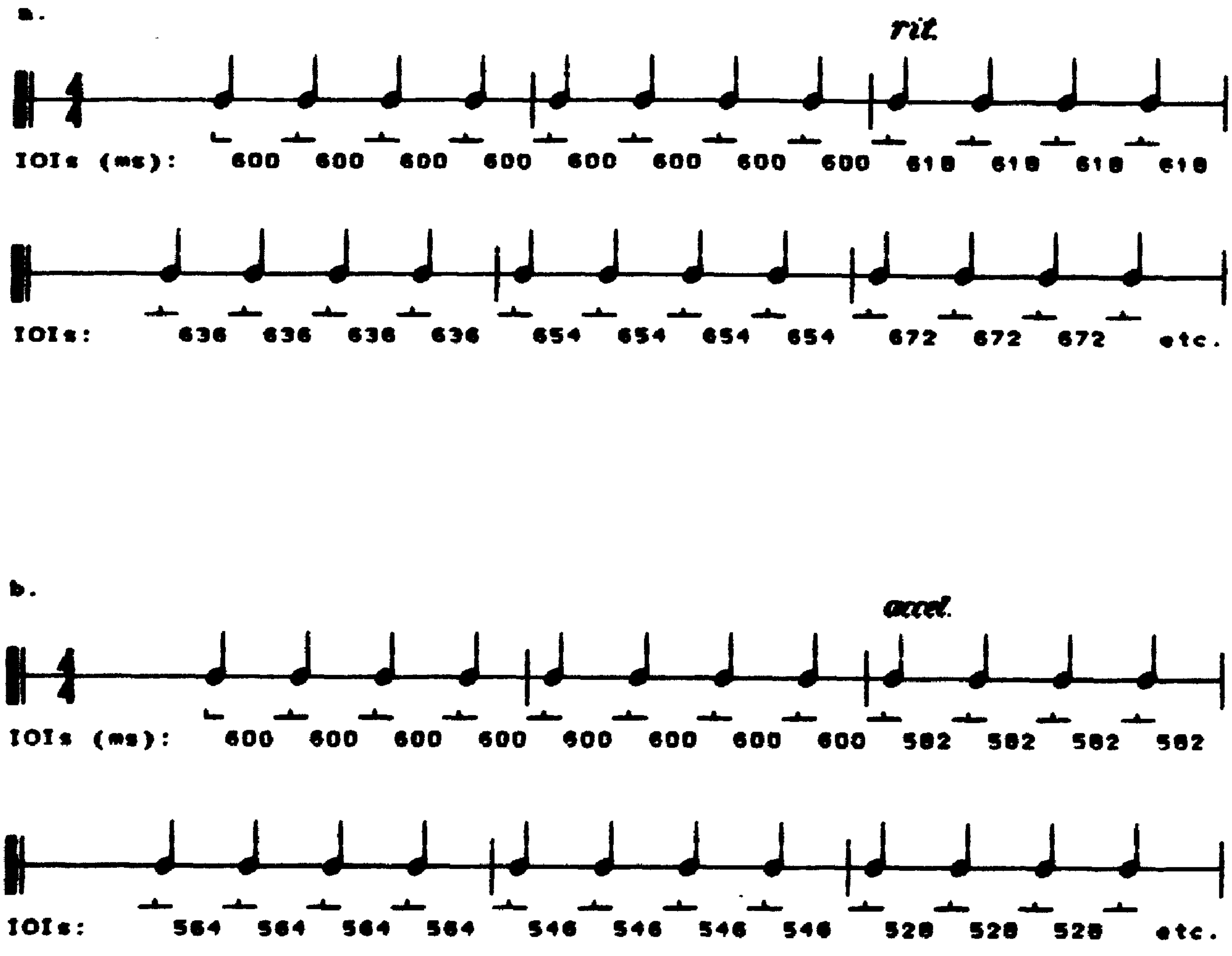


Figure 1.2. Two basic types of tempo modulation. Figure a shows a gradual deceleration: IOIs get ever longer and successive onsets get increasingly later relative to the initial isochronous rhythm. Figure b shows a gradual acceleration: IOIs get ever shorter and successive onsets get increasingly earlier relative to the initial isochronous rhythm.

example, who is concerned with coining a terminology sufficiently accommodating to be applicable to Renaissance and Twentieth-century as well as Common-Practice tonal music, offers a flexible definition:

> *Meter is that aspect of structure articulated as accent-delineated groupings within the attack (event) sequence, and the proportional interrelations of such groups at all levels.* . . . True metric structure is neither necessarily regular nor necessarily coincident with notated bar-lines . . . (Berry 1987, 318)

Cooper and Meyer are similarly forgiving concerning regularity:

> Meter is the measurement of the number of pulses between more or less regularly occurring accents. Therefore, in order for meter to exist, some of the pulses in a series must be accented--marked for consciousness--relative to others. (Cooper and Meyer 1960, 4)

On the other hand, theorists who have concentrated almost exclusively on the tonal repertoire are more insistent on stipulating regularity. Yeston, for instance, asserts that,

> The fundamental logical requirement for meter is therefore that there be a constant rate within a constant rate--at least two rates of events of which one is faster and another is slower. (Yeston 1976, 66)

Lerdahl and Jackendoff state:

> The term meter, after all, implies measuring--and it is difficult to measure something without a fixed interval or distance of measurement. Meter provides the means of such measurement in music; its function is to mark off the musical flow, insofar as possible, into equal time-spans. In short, metrical structure is inherently periodic. . . . Fundamental to the idea of meter is [also] the notion of periodic alteration of strong and weak beats; . . . For beats to be strong or weak there must exist a *metrical hierarchy*--two or more levels of beats. (Lerdahl and Jackendoff 1983, 19)

Thus periodicity on at least two levels is the *sine qua non* of meter for Yeston and Lerdahl and Jackendoff.

Both camps have their merits. Berry's and Cooper and Meyer's flexibility must be embraced if one is to allow so-called asymmetrical meters such as 7/8 or 5/4, as well as the type of additive meters, (e.g., 3 + 3 + 2/8) found in works, say, by Bartók or Stravinsky. On the other hand, Lerdahl and Jackendoff's point about the necessary regularity of a means of measurement seems well taken. Bearing the above considerations in mind, one might propose the following definition of meter:

- *Meter* is the abstract hierarchical arrangement of time intervals such that longer time intervals encompass whole numbers of shorter time intervals.

Time intervals that comprise the metrical hierarchy are abstracted from both IOIs and IAIs presented in the musical stimulus. At least one level of the hierarchy must comprise isochronous intervals. Other levels do not necessarily have to be isochronous, although at least one other level usually will be. In addition, in keeping with the temporal upper limit used to define rhythm above, no time interval in the metric hierarchy may exceed the limit of the psychological present (i.e., *circa* 8 seconds). Figure 1.4 illustrates three common metrical hierarchies in terms of both proportional IOIs and note-values.

Adoption of this definition still requires the identification of the locus of meter. To disentangle this question three interrelated types of meter should be recognized: *notated* meter and two types of *perceived* meter. Consideration of Figure 1.5 will serve to illustrate the distinction between the three types.

Firstly, the notated meter in Figure 1.5 is 3/4. This notational designation may consciously or unconsciously influence the performers' interpretation of the printed music. However, any conscientious performer will primarily have to respond to the notated

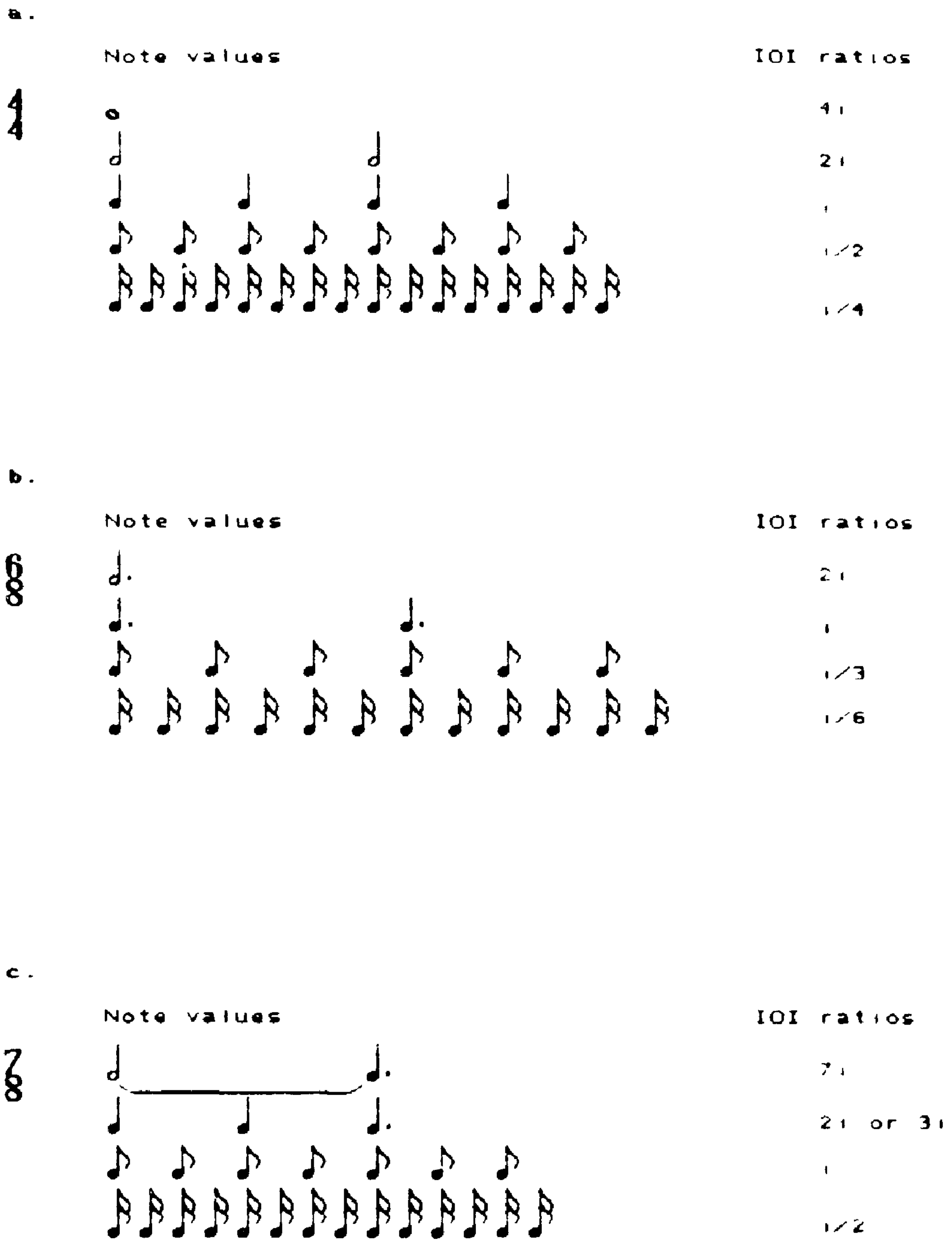


Figure 1.4. Metrical hierarchies in terms of note values and relative time intervals. Figure a shows a 4/4 meter with the quarter note as the beat-level time interval (*i*). Figure b shows a 6/8 meter with the dotted-quarter note as the beat-level time interval (*i*). Figure c shows a 7/8 meter with the eighth note as the beat-level time interval (*i*).

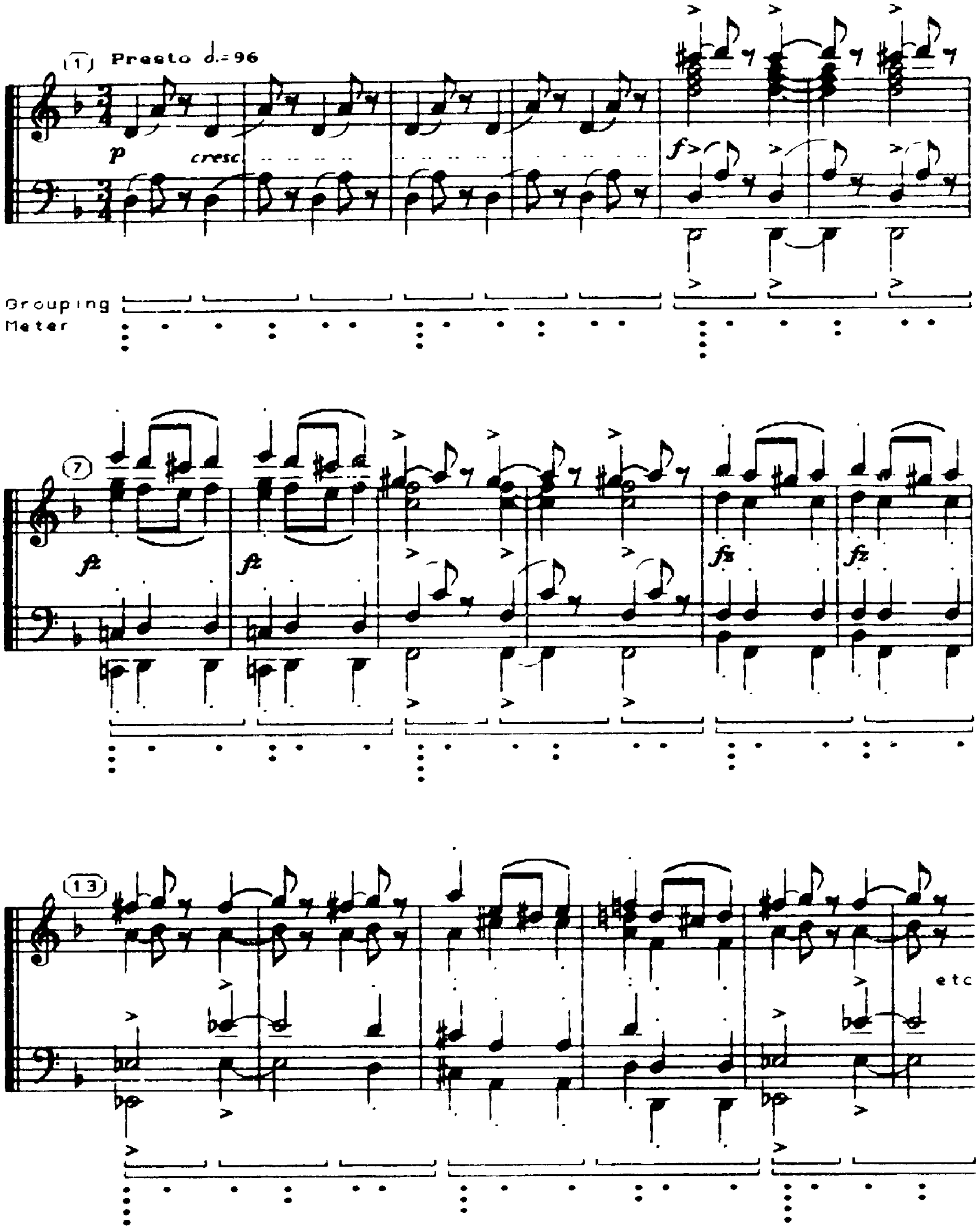



Figure 1.5. Dvořák, *Symphony No. 6 in D major*, Op. 60, III, bars 1–18. Figure shows orchestral reduction with grouping and metrical analyses *à la* Lerdahl and Jackendoff (1983).

stresses and articulation markings, so that the performance accents will suggest a meter that is alternating between 2/4 (e.g., bars 1-6) and 3/4 (e.g., bars 7-8).

The perceived meter is more complicated. A naïve listener might take the performed accents at face value and alter his/her metrical representation of this extract in synchrony with the accentuation as time unfolds. An alternative strategy might be to bring more abstract musical knowledge to bear in deciding on a representation of the meter. Abstract musical knowledge could consist of previous acquaintance with the work, or knowing that this piece is the scherzo of a symphony and that such movements are usually in 3/4. Jackendoff would argue that such knowledge cannot influence the note-by-note representation of meter since the cognitive processing apparatus responsible for arriving at a mental representation is "informationally encapsulated" from long-term memory (Jackendoff 1992, 67). Thus, in Jackendoff's view, even if the listener is consciously aware of Dvořák's ingenious metrical games, the listener's perceptual processor cannot make use of this knowledge. Having said this, however, Lerdahl and Jackendoff (1983) would presumably analyze a listener's *final* representation of the whole movement as 3/4. The strong inducements to 2/4 in bars 1 to 6 (and elsewhere) would be regarded as examples of grouping contradicting the overall metrical structure (metrical and grouping analyses *à la* Lerdahl and Jackendoff are included in Figure 1.5). Thus, the logical consequence of combining Lerdahl and Jackendoff's (1983) approach and Jackendoff's (1992) train of argument is that there must be at least two types of perceived meter, defined as follows:

- *Perceived real-time meter* is the changing representation of metrical structure that is constructed by the listener while the music is in progress.
- *Perceived abstract meter* is the overall representation of metrical structure that is constructed by amassing perceptual information in memory once the piece is completed.

Introspection suggests that the listener will strive to find one abstract meter that can account for an entire musical piece. Both the theoretical and practical portions of this dissertation will concentrate on perceived real-time meter, although perceived abstract meter will be addressed again in Chapter 3.

The final term to be defined here is *rhythmic categorization.* Once a metrical framework has been established, it is used by the listener to process time intervals that occur subsequently in a piece of music. Time intervals that are mildly non-conformant with the established metrical hierarchy, are categorized by the perceptual system as if they were conformant. Thus, for example, if the normative IBI is established as 500 ms (120 bpm), then IOIs such as 493, or 511 ms will probably be categorized as corresponding to the beat time interval. By the same token, intervals of, say, 243 or 261 ms will probably be categorized as duple subdivisions of the beat. Rhythmic categorization may therefore defined as follows:

- *Rhythmic Categorization* is a certain amount of perceptual leniency with inaccurately or expressively executed musical time intervals that do not strictly conform to the established metrical structure.

Thus, non-conformant time intervals are assimilated into the prevailing meter. This type of accommodation makes intuitive sense since many of the discrepancies between performed time intervals will be the result of human performer error, and will not be meaningful within the musical context. Obviously, however, if there are large inconsistencies between successive time intervals, say 500 ms followed by 578 ms, then the differences will be detected by the average listener. In this case, such a large discrepancy is likely to signify something deliberate and meaningful in the music, such as a ritardando, therefore we would not expect the perceptual system to ignore such phenomena.

The Nature of the Problem

Rhythm modulation and tempo modulation, as defined in the foregoing paragraphs, are two variants of the more generic concept of *temporal change*. The perception of rhythm and tempo modulation involves detection of a change in IOI from a previously established, normative time interval (normally the established beat). Rhythm and tempo modulation differ, however, in the way an established time interval is altered. Rhythm modulation maintains a steady beat rate across bars, but isochrony within bars is disrupted in an ever-more-pronounced fashion. On the other hand, tempo modulation, as employed in this study at least, manifests itself as a gradual change in the IOI across bars while maintaining isochrony within one bar.

Given the difference in time-span over which the two types of temporal change operate, it would be of interest to know if detection of rhythm modulation differs from the detection of tempo modulation across treatment conditions when various aspects of rhythm are varied. The principal problem addressed in this study may therefore be stated in general terms as follows: *How is the detection of rhythm modulation and tempo modulation affected by the rhythmic context within which these temporal changes are couched?*

The three experiments described in Chapters 4 and 5 examine the relative difficulty/ease with which temporal change is detected in different contexts. Using synthesized drum beats, Experiment 1 determines if, and under what conditions, one type of change is perceived more easily than the other. Experiment 1 also examines temporal change perception with both shortening and lengthening IOIs, and ascertains if such change is perceived differently depending on the starting tempo. Also using drum beats, Experiment 2 is primarily intended to investigate the possible effect of metrical location (or position in the bar) of temporal change on change perception. Experiment 3 examines the influence of beat subdivision on temporal change perception. Experiment 3 uses extracts of real piano music instead of drum beats, and so introduces a degree of ecological validity into the study. The factors listed below are manipulated in the three experiments:

1. TYPE. Two types of temporal change are examined: *rhythm* modulation and *tempo* modulation.

2. DIRECTION. Onsets can be displaced in two directions: the First Anomalous Onset (FAO) can occur *earlier* than expected or *later* than expected. Examples of *early* and *late* FAOs are given in Figure 1.6.

3. INITIAL BEAT RATE. In order to determine whether particular temporal-change-detection trends are pertinent to the entire musical

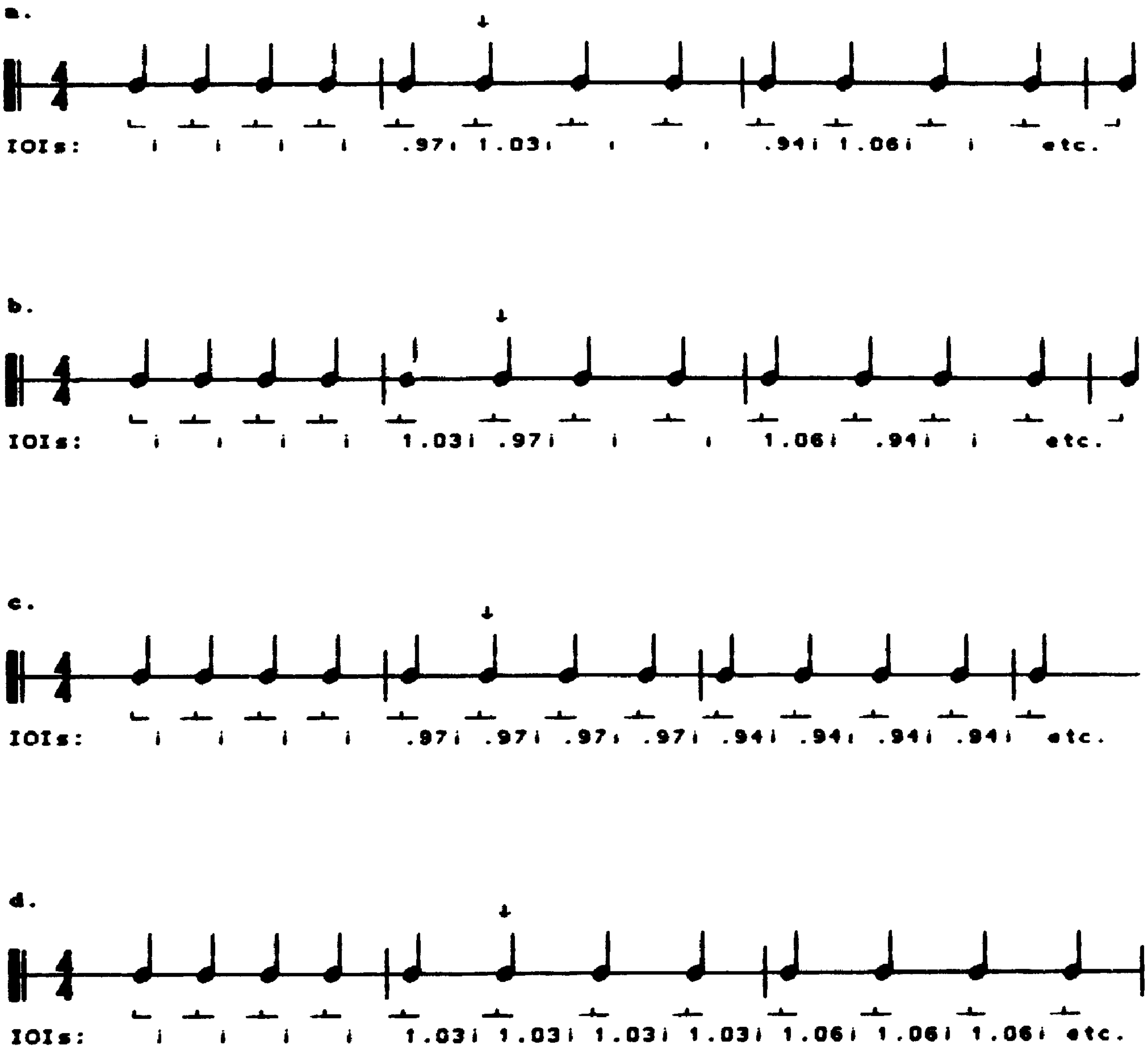


Figure 1.6. Early and late anomalous onsets in both rhythm modulation and tempo modulation. In each case the first anomalous onset (FAO) is marked with an arrow; IOIs are shown as proportions of the initial IOI, *i*. Figure a shows rhythm modulation with an early FAO. Figure b shows rhythm modulation with a late FAO. Figure c shows tempo modulation with an early FAO. Figure d shows tempo modulation with a late FAO.

tempo range, a number of different initial beat rates are employed across experiments: 45, 60, 75, 120, 150, 190 and 300 bpm.

4. LOCATION. Experiment 2 employs two LOCATION conditions, with the FAO on *beat 1*, or with the FAO on *beat 2*. In Experiments 1 and 3 LOCATION is held constant on *beat 1*.

5. PIECE. In order to test for the effect of beat subdivision on change detection, Experiment 3 uses four different extracts of keyboard music. Three of these pieces, those by Bach, Chopin and Hindemith, present sixteenth-note subdivision of the quarter-note beat, whilst a fourth piece by the author presents only the quarter-note beat. The levels of the PIECE variable, therefore, are: *Bach*, *Chopin*, *Hindemith* and *Author*.

6. VERSION. Finally, in order to test for reliability, each respondent is presented with both an *original* and a *replication* version of each stimulus condition. This replication allows one to determine whether learning or fatigue takes place over one experimental session.

Specific experimental hypotheses arise out of the perceptual model described in Chapter 3, and one must await a detailed delineation of this model before the hypotheses can be advanced in full. Nonetheless, it is possible at this stage briefly to outline certain predictions about the effects of manipulating the aforementioned variables. Generally, one overall prediction might be made: although rhythm and tempo modulation are two classes of temporal change, it is the principal hypothesis of this dissertation that they are perceived differently. The reasoning behind this hypothesis may be briefly summarized as follows: It is evident from Figure 1.6 that the change in note-to-note IOI-ratios in rhythm modulation becomes ever more marked as the stimulus progresses. For this reason, one might suppose that rhythm modulation is detected on a note-to-note basis, that is by comparing immediately adjacent IOIs in perception. On the other hand, in the case of tempo modulation, the actual change ratio from bar to bar (beat to beat) is quite small, for example 1.03/1 = 1.03, 1.06/1.03 = 1.029, 1.09/1.06 = 1.028 and so on. Therefore, if tempo modulation is noticeable, its detection must be due to a *cumulative* effect over several bars. Thus gradual tempo modulations, of the sort used in this experiment at least, will be detected by comparing the currently sounding tempo with some representation of the initial tempo in memory. From this reasoning, and from the results of the literature summarized in Chapter 2, one would expect different detection patterns for tempo modulation from those for rhythm modulation. This prediction leads to the principal hypothesis of this dissertation:

- Temporal changes that occur "suddenly," i.e., within the span of the psychological present, are detected in perception; temporal changes that occur "gradually," i.e., outside the span of the psychological present, are detected by comparing a memory trace with perception.

The different processes used to detect sudden changes (for example rhythm modulation) and gradual changes (for example tempo modulation) are predicted to result in statistically significant two-way interactions between the independent variable TYPE and other variables. Specifically, the following four subsidiary hypotheses may be identified. The first three are as follows:

1. There will be a statistically significant interaction between TYPE and DIRECTION
2. There will be a statistically significant interaction between TYPE and INITIAL BEAT RATE
3. There will be a statistically significant interaction between TYPE and LOCATION.

For reasons arising from the model outlined in Chapter 3, it is impossible to predict whether there will be a statistically significant interaction between TYPE and PIECE (subdivision) or whether PIECE as a main effect will be significant. Therefore, the fourth hypothesis is stated as follows:

4. There will either be a statistically significant main effect owing to PIECE alone, or there will be a statistically significant interaction between TYPE and PIECE.

These four hypotheses are tested in the experiments described in Chapters 4 and 5.

Importance of the Problem

The importance of investigating the detection of rhythm and tempo modulation is both theoretical and practical. Experimental work on meter and categorical perception has generally been informed by quite thorough theorizing. The main theories (Povel 1984; Clarke 1985; Jones and Boltz 1989; Parncutt 1994) will be reviewed in Chapter 2. However, as admirable as some of these attempts are, there has been a tendency to take a "bird's-eye view" of rhythm perception--as if most or all of the time intervals in a rhythm were available to the perceiver instantaneously and concurrently. Whilst one cannot deny that a certain amount of temporal integration must go on as perception is taking place in time (how otherwise could events be perceptually related?), the purpose of the theory proposed in this dissertation is to concentrate on the bottom-up, event-by-event processing of musical rhythms.

Much experimental research has already been carried out on tempo modulation perception. However, little or no attempt has been made to provide a theoretical account for the reliable trends that have been observed. Rather, researchers seem to have been daunted by the multiplicity of tempo-related behaviours, and have simply acceded to Wang's view that "the nature of tempo perception is evidently quite complicated." (Wang 1983, 55). This dissertation does not intend to deny the complex nature of tempo perception. However, it is hoped that by placing the relevant experimental results within a wider framework of time-interval perception in music, some tentative explanatory theory can be put forward.

A possible practical value of this study may be that, in contributing to a greater understanding of the human processing of temporal information in music, it can inform music-making activities for music educators and performers. In music education, novice performers often find the temporal aspects of music challenging. If educators know why a student cannot hear that he/she is playing a rhythm incorrectly, or why he/she is straying wildly from a steady tempo, this knowledge offers the educators more efficacious tools for correcting these faults in an informed manner before bad habits become ingrained. In addition, knowing which aspects of time-interval perception are more difficult than others, it is possible that educators may better arrange the order of different types of rhythmic exercises in a pedagogical sequence (for an attempt at ordering, see Gordon 1976).

In the case of performance, expert players and singers frequently employ various types of rhythm modulation (agogic accents, *inégalités*, swing) and tempo modulation (rubato, ritardando) to breathe expressive life into their performances. A greater appreciation of the extent to which these expressive devices are perceived may provide new information about the psychological processes underlying successful performances. In particular, experimentation investigating the parameters of rhythm and tempo modulation may help to identify specific contexts in which these two types of temporal change are most likely to be perceived in musical performance.

Organization of the Dissertation

The perception of temporal change within a metric context will be examined from both a theoretical and an experimental standpoint. The theoretical portion of this dissertation comprises mainly Chapters 2 and 3. Chapter 2 provides a review of the pertinent literature on the perception of musical rhythm. Firstly, general ideas from "time psychology" and memory research are introduced and their applicability to music is discussed. This discussion leads into consideration of such topics as tempo modulation perception, beat perception, meter and categorical perception of rhythm. The general findings of Chapter 2 are summarized, and then Chapter 3 is set the task of accounting for at least some of these results through a theoretical model. The model generates hypotheses concerning the manner in which short time intervals in music are perceived and cognized. Although the model is fairly general, several specific predictions arise from the interactions of different stages within the model. In order to test the model against introspection, several rhythmic stimuli are considered as they "pass through" different stages of the processing system.

Whatever intuitive appeal explanations in Chapter 3 might have, real-life empirical testing of the theory is of course necessary. The three experiments described in Chapters 4 and 5 are designed to test some of the hypotheses arising from this model. Experiments 1 and 2 use fairly simple unpitched drum stimuli whose rhythmic characteristics are tightly controlled. This temporal control is exercised in the interests of internal validity. However, one unfortunate sacrifice that is made whenever internal validity is sought is that stimuli must be simplified compared to what one would encounter in real-life musical situations. The isochronous drum rhythms employed in experiments 1 and 2 are a far cry from the rhythmic subtleties of a typical musical performance. However, this study is experimental, and, in the experimental tradition, it is a tenet of this dissertation that the better part of wisdom lies in working from the ground up. It is only by first understanding the perception of rhythm and tempo modulation of raw beats, that one can then progress to understanding how temporal change is perceived in more complex rhythms.

The above arguments notwithstanding, Experiment 3 essays to imbue the study with a little more ecological validity. Extracts of real piano music are used to see how the same temporal changes are perceived in the midst of a melodic/harmonic phrase, rather than in simple drum beats. It should be added, however, that the stimuli used in Experiment 3 still lack certain attributes found in real musical performance. Most notable of these, bearing in mind the focus of this dissertation, is rhythmic expression, viz rubato and agogics. In other words, apart from the rhythm and tempo modulation introduced as control variables, the piano-piece excerpts are still metronomic (controlled by a computer sequencer). For this reason, they still maintain a modicum of artificiality. However, the alternative, adding controlled rhythm or tempo modulation *on top of* other expressive temporal changes, seemed less desirable. Since the purpose of the experiments was to detect temporal change, in such circumstances it would have been impossible to direct subjects' attention exclusively to the controlled temporal changes. For this reason, one could not know whether a given response datum reflected the perception of rhythm or tempo modulation, or other expressive deviations.

The final chapter, Chapter 6, draws the results of the three experiments together. Experimental findings are interpreted in the light of both the experimental literature and the proposed model. Possible confounds are addressed and, to bring the dissertation to a close, suggestions for future research are made.

CHAPTER 2

REVIEW OF THE LITERATURE

As in many areas of research, the list of citable literature pertaining to the psychology of rhythm, meter and tempo is long and varied. In order to reduce the sheer volume of work that could reasonably be surveyed, this chapter will focus on those studies that have a direct bearing on rhythm and tempo modulation perception. However, so that individual experimental studies may be put in a general context, the first part of this literature review will provide a conspectus of several aspects of time perception in general. The chapter is therefore divided into three sections:

1. *Time Perception.* Theoretical ideas from research in the psychology of time set the scene for time-interval and temporal change perception in music. The notion of the psychological present is explored as the limit of one span of attention. One possible mechanism that may contribute to the psychological present in the auditory modality is "echoic memory," which is introduced next. Consideration of echoic memory leads naturally into other, less evanescent forms of auditory memory. Finally in this section, a different view of attention is examined, namely that attention is periodic and mediated by internal oscillators.

2. *Tempo Perception.* The second section surveys the experimental literature relevant to tempo modulation perception. This survey begins by examining the range within which listeners preferentially choose a beat-rate. Having established the preferred tempo range, the section goes on to identify under what conditions tempo modulation is optimally perceived.

3. *Duration, Beat and Meter Perception.* The third section discusses the literature that pertains to rhythm modulation perception. The results of psychophysical experiments on duration discrimination are contrasted with studies investigating abstract cognitive representation of musical time intervals. The importance of beat extraction for the cognitive representation of rhythm is discussed, and this in turn leads to consideration of meter and the categorical perception of rhythm.

Whilst this literature review attempts to cover all the major research that has some relevance to the theoretical and experimental portions of this dissertation, it is not exhaustive in all areas and some subjects are only cursorily represented. Firstly, the experimental literature on grouping, a concept which seems central to many people's notion of rhythm, is not examined. The reason for this apparent disregard is that grouping (and therefore Cooper and Meyer's theory) is only of passing interest to the model of time-interval perception forwarded in this dissertation. By extension then, much of the experimental work on patterning arising from parameters such as pitch, dynamics or timbre--auditory stream segregation for example--is not included. Secondly, possible links between rhythmic ability and physical movement are outside the scope of this dissertation. Studies concerning psycho-motor or performance-oriented behaviour are mentioned only insofar as they inform the discussion concerning time-interval or meter perception. Thirdly, an established corpus of literature on the affective and/or motional quality of rhythm is omitted (e.g. Gabrielsson 1973). All three of the above bodies of research assemble quite legitimately under the banner of rhythm, and their omission is not intended as an implied rejection of their purport or methodology. Rather, they represent such a diversity of approaches and opinions that their integration into a single theory of rhythm must await future research.

Time Perception

Block (1990) identifies three general aspects of psychological time that have consistently received attention from researchers: (1) succession (perception and recognition of sequential order), (2) duration (perception and categorization of durations in certain contexts) and (3) temporal perspective (individual experiences of past, present and future). All three of these aspects of psychological time can conceivably have some effect on music perception and cognition. However, this dissertation, and therefore this chapter, focus on item number 2, duration perception, which covers the three central concepts of musical time: rhythm, meter and tempo.

The Psychological Present

A number of models of time perception regard temporal experience as a sampling process in which the flow of stimulus information is divided up into either discrete or continuous chunks. These models are often called *perceptual moment models* as they imply the integration of an extended span of physical time into what is perceived as a single instant (Block 1990). The most venerable perceptual moment model, and the one most invoked in discussions on rhythm, is that of the *psychological present*, first denoted by William James as the *specious present* (James 1950).

For James, the specious present was a continually shifting perceptual window (or travelling moment) through which one could integrate events that are physically separated in time. As James himself put it "The practically cognized present is not a knife-edge, but a saddle-back, with a certain breadth of its own, on which we sit perched and from which we look in two directions into time" (James 1950, 609). The concept of the specious present has proved enduring, possibly because it is phenomenologically appealing and theoretically useful. However, since James' original formulation, the specious or psychological present (the latter being the currently preferred term) has undergone some necessary revisions. Most importantly, perhaps, the idea of a continuous, shifting window is now regarded as problematic since such an arrangement would lead to multiple experiences of the same (physical) event (Mabbott 1951). As Fraisse (1963) points out, we do not at one instant hear the "tick-tock" of a clock, and at the next its "tock-tick," but rather group its ticks in pairs consisting of perceptually accented and unaccented beats. This observation challenges James' conception of the specious or psychological present as a sort of perceptual gang-plank, which, as it assigns one perceptual event to a cabin in long-term memory, admits a new stimulus event from the dockside of the physical world. For this reason, the psychological present is now regarded as a discrete moment, a self-contained chunk of attention. Accordingly, a revised definition of the psychological present is offered by Michon: ". . . the time interval in which sensory information, internal processing, and concurrent behavior appear to be integrated within the same span of attention." (Michon 1978, 89).

In musical rhythm the notion of the psychological present seems particularly useful; the fact that listeners are able to relate consecutive notes or chords to each other, to group them, suggests that they routinely make use of some type of temporal integration. Some agreement is evident among writers concerning estimates of the typical time span of this integration. Fraisse (1984) states that it seems to have an upper limit of about 5 seconds, but is more typically nearer to 2-3 s in time span. Michon (1978) agrees with the average value of 2-3 s but places the upper limit at 7-8 s. Pöppel (1972) has similarly estimated the maximum duration of the psychological present to lie between 4 and 7 s. The absolute lower limit of the psychological present may be set at the threshold of event succession perception (*circa* 50 ms), although Michon (1978) notes that in many cases events must be separated by at least 150-250 ms to be perceived as individual entities. Whatever the extreme limits of the psychological present are, Michon (1978) adds that temporal structure or grouping is most apparent when events are separated by between 250 and 900 ms.

However, to talk of the psychological present in terms of its typical temporal spans can lead one to imagine that it is a single cognitive or neurological mechanism. Rather, it should be stressed, the psychological present is a *phenomenal byproduct* of perhaps multiple perceptual and cognitive processes that all attempt to find pattern in and impose order on a wide variety of incoming stimuli. As such, the time span of one psychological present may be expected to vary greatly depending on the sensory modality, the rate of succession, and the complexity of stimulus events. Such a diversity of observed temporal spans is evident in the literature on the hypothetical "echoic memory," one of the likely components of the psychological present in the auditory modality. The parameters of echoic memory will be examined next.

Echoic Memory and Short-Term Memory

The echoic memory (a term originally coined by Neisser [1967]) is presumed to hold a single, discrete chunk of the auditory signal. Echoic memory preserves most of the stimulus' acoustical detail intact, yet this detail decays inexorably over several seconds. As such, echoic memory provides precategorical storage of stimulus events and allows integration and comparison of events over a short time-span. Echoic memory has also been called "precategorical acoustic storage" (Crowder and Morton 1969), or, more recently, "phonological loop" (Baddeley 1990).

The duration of echoic memory has been estimated using a number of different experimental approaches.[1] Work on dichotic listening has

---

[1] Although they are counted as manifestations of echoic memory by some reviewers (e.g., Cowan 1984), the psychophysical phenomena of auditory persistence, (e.g., Efron 1970), amplitude-duration integration

suggested duration estimates for echoic memory of 1.5 s (Treisman 1964), approximately 2 to 4 s (Bryden 1971), and no more than 5.3 s (Glucksberg and Cowen 1970). Studies using a stimulus recognition paradigm have given estimates of circa 2 s for vocalized digits and letters (Darwin, Turvey and Crowder 1972), whilst Treisman and Rostron (1972) noted a near chance recognition of tones after 1.6 s. In a periodicity detection study, Guttman and Julesz (1963) found that subjects retained spectral information about a white-noise stimulus for up to 2 s, whilst Kubovy and Howard (1976) estimated the half-life of the echoic store to be around 1 second. Using a two-stimulus-comparison paradigm, Crowder (1982) found that vowel comparison deteriorated markedly after 3 s, whilst Lu, Williamson and Kaufman (1992) found that memory for tone loudness had a decay half-life of *circa* 2 s, and reached its asymptote around 5 s.

Yet another batch of studies has suggested that some acoustical detail is left to the listener after much longer delays: in a review of the literature on modality and suffix effects (e.g. Crowder and Morton 1969, Morton, Crowder and Prussin 1971), Cowan (1984) suggests that auditory memory may last up to 20 s. In addition, Wickelgren (1969) noted a continued decline in comparison performance when tones were separated by intervals greater than 24 s.

How might one account for such a wide variety of estimates for the duration of echoic memory? Firstly, as Kärnbach (1994) has noted, when considering a limited-span store such as echoic memory, at least

---

(e.g., Stevens and Hall 1966) and detection masking (e.g., Zwislocki 1972) are here considered to result from more peripheral neural mechanisms than those responsible for echoic memory.

two durational parameters should be identified: (1) storage capacity and (2) decay time. The storage capacity of echoic memory would be the maximum storable length of one "chunk" of stimulus, whereas the decay time would be the time the trace takes to deteriorate beyond recognition. In many studies on echoic memory, experimenters simply assume that they are examining decay time, and no mention of storage capacity is made. However, if one envisages echoic memory as possessing an input buffer that can integrate, say, up to 2 seconds of stimulus signal, once that buffer is full, its contents would either have to be shunted elsewhere to make room for new incoming stimulus information, or be jettisoned. Both the above literature review and introspection suggest that auditory traces of only a few seconds old are not entirely jettisoned, implying that they are relocated. If, as part of the relocation process, acoustical fidelity is lost, then this process would be indistinguishable from trace decay. In short, the above estimates of the duration of echoic memory may reflect its storage capacity, its decay time or both.

A second reason for the variety of durational estimates for echoic memory is the diversity of experimental tasks asked of the subjects in the above studies. For example, if a subject has to compare the timbres of complex sounds (white noise for example), this type of comparison task will suffer much more quickly from signal decay in echoic memory than will a task where a subject is asked to compare the pitches of two successive sine-waves. In addition, the nature of a number of the experimental tasks in the above literature (for example word recall) suggest that other more abstract types of memory may influence subjects' performance.

Generally, in fact, longer recall times of more than about 3 s are reminiscent of what has been called short-term memory (Miller 1956). In the classic information-processing model of perception, perceptual input is regarded as a three-stage serial process involving a "sensory register," "short-term memory" (or STM) where only a limited number of partly categorized elements can be manipulated, and an almost limitless "long-term memory" whose contents are more abstract (Michon 1985). In such a model, echoic memory would lie at the interface between the sensory register and short-term memory. Yet, once one strays into the area of memory, one has, strictly speaking, left the domain of the psychological present and temporal integration in immediate perception. As Michon notes, the psychological present and STM ". . . are essentially different: the 20-30 sec estimates [for STM] are obtained with a limited *static* load on memory, whereas the loads on the present are dynamic . . ." (Michon 1978, 95).

However, as the literature reviewed above suggests, the conceptual distinction between echoic memory and STM is frequently blurred. One oft-muted difference is that, whereas echoic memory holds information in a precategorical state, the contents of STM seem to be a mixture of raw signal data and more abstract categorical information about the stimulus (Cowan 1984). This mixture of raw and abstract information is evident in some of the experiments performed on STM for musical stimuli. For example, Dowling and Bartlett (1981) and Dewitt and Crowder (1986) found that whilst interval structure was important for recalling melodies in long-term memory, the more primitive factor "contour" was more influential for recall in STM. However, Cuddy and Lyons (1981) found evidence that in spotting mistakes in two consecutive transpositions of a standard melody, subjects had a lower percentage-correct score for modulating and nondiatonic melodies than they did for diatonic melodies. Similarly, Croonen and Kop (1989) submitted that intervallic information might be extracted from sequences in STM, provided strong cues for a particular key centre are presented with the melody. The results of both Cuddy and Lyons (1981) and Croonen and Kop (1989) suggest that some degree of abstraction occurs in auditory STM.[2]

One noteworthy model of auditory short-term memory effectively combines echoic memory with STM (Frick 1988). Frick proposes that auditory short-term memory may be likened to a variable-speed tape recorder: if storage of fine-grained acoustic detail is required, the tape must run fast to capture the desired fidelity, and hence the maximum duration (capacity) that can be recorded is correspondingly reduced. Alternatively, if a premium is placed on recording a long stimulus, the tape can run slowly, but with a corresponding drop in trace fidelity. Evidently, the tape recorder analogy suggests that the contents of auditory short-term memory are raw acoustical data. Frick therefore posits that this unprocessed information is parsed and categorized just prior to recall. What is not made clear in Frick's model, however, is how long a trace may be stored and what happens to it

---

[2] The above references are intended merely as examples of studies that offer evidence for a mixture of precategorical and abstract information in STM for musical stimuli. For a more comprehensive review see Berz (1995).

once a new stimulus is presented to the listener. In other words, although Frick addresses the question of storage capacity, he neglects the issue of trace decay. Presumably, if recorded acoustical data is to be used at all, abstract, parsed information must be extracted from it before it has decayed beyond usefulness or been replaced by a new stimulus. These issues will be addressed by a modified version of Frick's model in Chapter 3.

For the purposes of this dissertation, it will be assumed that the storage capacity of echoic memory is typically 1-2 s and no longer than 3 seconds. Much of the literature reported above converges on this approximate duration range. Longer values of 20 to 30 s reflect the decay time of echoic memory, rather than its storage capacity. It is here suggested that auditory STM is simply the accessing (the playing back) of this decaying auditory trace, combined with the more abstract processing operations that are performed on that trace (e.g., the superposition of a cognitive tonal scheme on a set of pitches). In this way, temporal integration in perception (the psychological present) is intimately tied up with immediate recall of auditory events (the "psychological past").[3] Both are evidently vital to the ongoing process of integrating and organizing temporally separated auditory events.

---

[3] Coincidentally, a more recent memory paradigm, namely *proceduralism*, supposes that memories are simply the by-product of perception and cognition, the useful residue of the combined perceptual and cognitive processes that come to bear on a stimulus (Crowder 1993). In the proceduralist viewpoint, however, there is no need for discrete storage-site type memories dedicated to specific sorts of stimuli or to specific sorts of tasks.

Attentional Periodicity and Internal Oscillators

The psychological present was characterized as a single span of attention in which stimulus events are integrated. Although the previous paragraphs have examined the typical duration of the psychological present and its probable component, echoic memory, nothing has been said about how attention might be focused over time. A temporal theory of attention has been developed by Jones and associates (Jones 1976, 1990, 1992; Jones and Boltz 1989; Jones and Yee 1993), and this theory will be described in the following paragraphs.

Jones (1976) first proposed her theory in the wake of an important experiment by Handel (1973). Handel's study had shown that simple auditory sequences are processed much more easily if they can be divided into groups of a constant time interval. To explain Handel's findings, Jones posited that the listener's attention tunes into the periodicities present in the first few seconds of a stimulus, and extrapolates these time intervals to predict when events, or groups of events, will occur in the future. Thus, if two events e1 and e2 are separated by a time interval *i*, then they will generate the expectation of a third event (e3) at i after e2 (see Figure 2.1). If *i* is duplicated between e2 and e3, it is strengthened as a plausible time interval for the prediction of yet more future events (e4 etc). Furthermore, the interval between e1 and e3 (*2i*) now sets up its own expectation and predicts the occurrence of e5. Thus this projection of time intervals into the future works on two (or more) levels, that is, it is *hierarchical*. Experimental evidence that regularly spaced

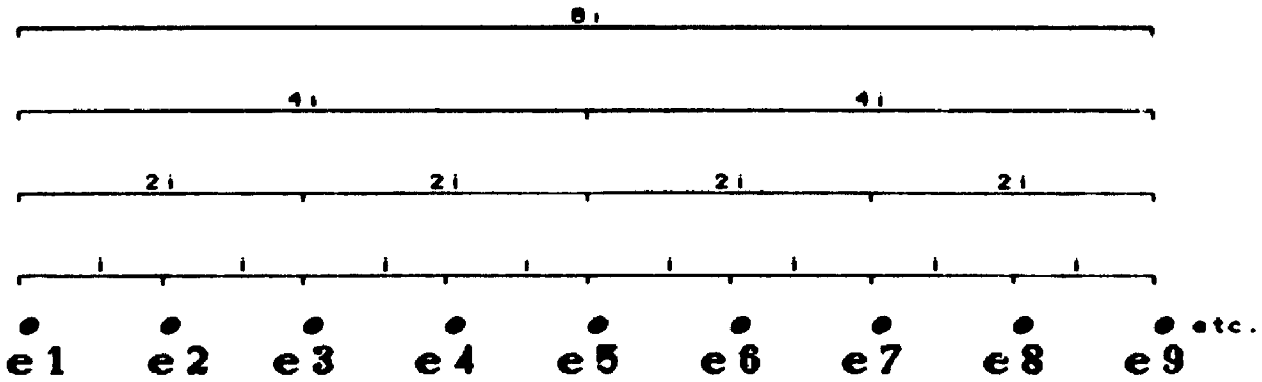


Figure 2.1. Binary hierarchy of time intervals, as envisaged by Jones (1976). Noteheads represent the onset of an auditory event (e1, e2, etc.); time intervals between events (IOIs) are represented as multiples of i. Time intervals at one hierarchical level are nested within time intervals twice as long at the next hierarchical level up. This hierarchy is therefore recursive since time intervals on adjacent levels are related by a constant factor of 2 (C_t = 2).

(isochronous) events on at least two hierarchical levels facilitate perceptual processing is offered by Essens (1986).

Jones and Boltz (1989) hold that this periodic focusing of attention plays an important role not only in the perception of musical sounds but also in the perception of non-musical sounds such as speech. Of course, for sounds in nature, where events often happen at seemingly random time points, such expectations are frequently confounded. However, the theory makes sense from the point of view of cognitive economy, for, if on the evidence of previous events a future event is likely to occur at a certain point in time, then it is sensible to save one's attentional effort for that moment. Notice also that, rather than focusing of absolute time intervals that may or may not be germane to time perception, Jones' theory emphasizes the role of *relative* timing. Events are processed more reliably if they occur at some multiple of the initial time interval *i* after e1. In theory at least, *i* can be any absolute time interval.

An expansion and refinement of Jones' (1976) theory may be found in Jones and Boltz (1989). These authors discuss the relevance of hierarchy at some length. Hierarchical structures involve the embedding of a set of equal time intervals at one level (T_n) into a time interval at the next level up (T_{n+1}). So, returning to Figure 2.1, the time interval at level T_n is *i* and the time interval at level T_{n+1} is *2i*. The relationship between the time intervals at T_n and those at T_{n+1} is constant and Jones and Boltz denote this constant as a ratio C_t. Thus, in Figure 2.1, $C_t = 2$ since $T_{n+1} = 2\ (T_n)$. In addition, Jones and Boltz regard recursiveness as a necessary feature of truly hierarchical structures. A recursive hierarchy would be one where the same constant ratio relates any pair of adjacent levels. Thus, for example, in a recursive hierarchy, if $T_{r+1} = 2 (T_r)$, then $T_{r+2} = 2 (T_{r+1})$ and so on. Figure 2.1 is an example of a binary recursive hierarchy.

However, it is Jones and associates' very insistence on recursiveness that can be problematic. The stipulation that a constant C_r should interrelate all adjacent levels of the hierarchy is unfortunate. If the theory is intended to model musicians' notions of meter and hypermeter (and Jones and Boltz's use of musical stimuli suggests that it is), then this stipulation must be relaxed. There are many commonly used meters that are not recursive, for example 3/4 which divides time spans of one bar (dotted half-note) into three beats (quarters), but then each of these beats is itself subdivided into *two* eighth-notes.

A second issue that has to be addressed with regard to Jones' theory is how the perceptual system might actually tune into the periodicities inherent in the stimulus. The most favoured explanation is that neural oscillators in the listener's brain *entrain* to the periods that elapse between successive stimulus events (Gross and Daan 1985). Entrainment is a concept originally developed in physics to describe the way that two mechanically coupled clocks gradually shift periodicity so as to tick in synchrony with each other. Although the hypothetical oscillator involved in time perception is not mechanical, the notion that neural "ticking" adapts its period to be synchronous with an external periodicity is evidently close to the original idea of entrainment. Although direct observation of entrainable neural oscillators is at present lacking, entrainment has been proposed as the way a wide variety of organisms adapt to periodic changes in the environment (Gross and Daan 1985).[4] In addition, a computer model of entrainment to short time intervals is provided by Large et al (1993).

By way of summary, it should be pointed out that this section on time psychology has discussed the processing of two types of time interval. Firstly, in talking about the durational ranges implied by the psychological present and echoic memory it has referred to *absolute* time intervals. Although the duration/decay time of these hypothetical mechanisms is never rigidly prescribed, there are certain upper and lower bounds beyond which events simply last too long or sound too rapidly to allow integration in one span of attention. Secondly, in talking about attentional periodicity, this section has referred to *relative* time. According to Jones' theory, a series of time intervals is coherent provided that they form simple integer multiples of the shortest time interval present in the stimulus. One of the tenets of this dissertation is that the complementary concepts of absolute and relative time both tell a truth about the nature of time-interval perception. The manner in which both absolute and relative time-interval values come to bear in music perception will be fleshed out in the following two sections on tempo and meter. Tempo perception, especially the preferred tempo range, may be seen as a manifestation of

---

[4] Rather, one should say that observation of neural entrainment to *rhythmic* durations is at present lacking. The frequency theory of pitch perception, for instance, posits that the firing rate of certain auditory neurons matches (entrains to) the frequency of the incoming stimulus, and a number of experimental studies support this theory (see Moore 1982, ch. 4, for a review).

absolute time intervals in music perception. Conversely, meter may be regarded as a musical embodiment of relative time intervals.

Tempo

The experimental literature on the perception of musical tempo may be broadly divided into two problem areas: (1) that concerned with the preferred beat tempo range, and (2) that concerned with the accuracy of tempo modulation perception. Since tempo modulation perception has been examined experimentally over much of the tempo range, a review of problem area (1) may serve as an introduction to problem area (2).

Preferred Tempo Range

Work on preferred or natural beat tempo range ranks as some of the most venerable research in music perception. Early experimenters were interested in the so-called "indifference interval," or the time interval between successive events that makes those events sound neither too slow nor too fast. Using simple metronomic stimuli, Wundt (1911) found an average indifference interval of 600 ms. This estimate is matched by the average "preferred tempo" of 100 to 120 bpm (500-600 ms) observed in Wallin's (1911) study, which had subjects compare tempo pairs until they homed in on a preferred tempo. However, Wallin also noted a large inter-subject variance with minimum and maximum preferred tempi as far apart as 40 and 208 bpm (1500 and 288.5 ms). The idea of a preferred time-interval of 600 ms between sound events is bolstered in other studies performed by Fraisse (1963). When asked to reproduce temporal intervals, subjects tended to overestimate ones shorter than

500-600 ms and underestimate ones longer than this. As a consequence of this observation, and because Fraisse also noted that prospective judgment of intervals is most accurate around the same time-interval range, the 500-600 ms range has been identified as the indifference interval.

Another approach favoured by researchers from the 1930s onwards was to record subjects' "spontaneous tempo" or "personal tempo." Spontaneous or personal tempo is the rate at which an individual will spontaneously tap his/her foot or finger in the absence of any cuing stimulus. Various aspects of personal tempo were examined by such researchers as Frischeisen-Köhler (1933) who proposed that this phenomenon was hereditary, or Rimoldi (1951) who sought to demonstrate a correspondence between the natural motoric rate of opposite limbs. In contrast, however, Harrison (1941) found "no indication of a unitary speed trait or factor which is characteristic of the voluntary spontaneous movements of an individual." (Harrison 1941, 373). Rather, the spontaneous and the maximum tempo of individuals performing various motoric tasks depended on the nature of the task, and Harrison found little inter-task correlation for each individual subject. More positive results were obtained by Fraisse and his co-workers in Paris (Fraisse, Pichot and Clairouin 1949; cited in Fraisse 1982). For tapping, Fraisse et al found that the preferred time-interval varied between 380 and 880 ms, although "one can assert that a duration of 600 msec is the most representative." (Fraisse 1982, 153).

Support for a slightly longer preferred time-interval range is found in Brown (1979). In reviewing some of the literature on tempo, Brown notes that a normal or preferred time interval is to be found around 700 ms but may drop as low as 100 ms. He goes on to add that "the 750 msec. area in both Time Psychology and Movement is inescapable. In addition, its metronomic equivalent of *circa* MM=80 . . . has long been considered the 'just' or 'normal' tempo." (Brown 1979, 23). However, accompanying Brown's 1979 paper is an empirical longitudinal study that had subjects play musical extracts under different conditions (Brown 1981). This second study seems to imply that whilst individual tempo preferences are remarkably consistent, various factors such as instrument on which a piece is performed, time of day and modality (major versus minor) give rise to a large inter-individual variability. Interestingly, a more recent repeated-measures study concerned with the consistency of listener (rather than performer) preference found that a subject's musical background and the style of a composition do not significantly affect tempo preference, but the initial speed of the stimulus music does (Lapidaki and Webster 1991). In other words, once listeners are presented with a certain tempo they exhibit preference inertia, a reluctance to change what is given.

Two studies by Handel and associates, concerned with the perception and reproduction of polyrhythms, provide further evidence for a wider preferred tempo range than the 500-600 ms span suggested by Fraisse (Handel and Lawson 1983; Handel and Oshinsky 1981). In both of these experiments subjects were asked to tap along with polyrhythms in any way they wished. The polyrhythms consisted of two or three non-integrally related pulse-trains (e.g., 2 against 5 [2x5] or 2 against 3 against 7 [2x3x7]) that were manipulated in terms of loudness, pitch and pattern-repetition rate. When pattern-repetition rate affected subjects' responses, it tended to produce a moderating effect: faster pulse trains were tapped at slower tempi and slower pulse trains at faster tempi. For all the polyrhythms presented in Handel and Oshinsky, the between-tap intervals extended between 100 and 1000 ms with a concentration around 400 ms (150 bpm). Similar time-interval ranges seemed to apply in the Handel and Lawson experiments although the introduction of three-component pulse trains made these results harder to interpret.

One ingenious paper studied undergraduates' preferred tempo for nineteen popular songs in both perception and imagination (Halpern 1988). For each song, subjects either adjusted the speed of a computerized performance until they attained optimum tempo, or they calibrated a metronome to match the tempo of an imagined performance in their heads. Preferred perceived tempi had a mean of 99.2 bpm (604.8 ms) and preferred imagined tempi had a mean of 97.6 bpm (614.8). In addition, for all songs, perceived and imagined tempi were highly positively correlated, although Halpern found that the minimum and maximum imagined tempi for each song varied quite widely, from 65 to 164 bpm (923 to 366 ms).

Three other related studies reveal a broad yet consistent preferred-tempo range. Duke (1989b) and Duke, Geringer and Madsen (1991) asked subjects to tap along to a stimulus pulse-train at tempi ranging between 40 and 200 bpm. Duke (1989b) found subjects tapped in synchrony only with those stimuli between 60 and 120 bpm, and that they subdivided or grouped together beats at tempi respectively below or above this tempo range. Similarly, Duke et al (1991) found that college music majors tapped in synchrony only with those stimuli between 70 and 120 bpm, although other less musically trained subject groups were less discriminating. For Duke (1989b) the modal tapping tempo was 80 bpm, and for Duke et al (1991) it was 90 bpm. Geringer, Duke and Madsen (1992) used a polytimbral isochronous sequence that modulated gradually over the entire musical tempo range. The stimulus presented three drum-machine sounds with the rhythmic ratio 4:2:1; subjects had to say which timbre carried the beat at any given time. These authors found that subjects tended to assign the beat to whichever timbre was pulsing in the 60 to 120 bpm range, with a median preferred tempo of 90 bpm.

Generally then, despite much inter-individual variance and the influence of such factors as musical style or mood, experiments suggest a broad preferred-beat-rate range between about 60 and 150 bpm (1000 to 400 ms) with an average around 90-100 bpm (670-600 ms).

Perception of Tempo Modulation

Many experiments on tempo modulation perception have tried to establish which stimulus-related variables expedite detection of beat-rate change. Some of the commonly examined factors include direction of tempo modulation (deceleration versus acceleration), initial tempo, the degree and rate of tempo modulation, and the presence of rhythmic/metric features such as beat subdivision and grouping. The ensuing review will group studies by independent variables.

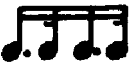
Kuhn (1974) had professional musicians listen to metronome pulses that gradually changed tempo from a predetermined point in the stimulus pulse-train. Subjects had to indicate when the tempo changed and how (increase, decrease or same), thus the dependent variables were reaction time and correctness of response. Four initial tempi were presented: 60, 90, 120 and 150 bpm (initial IOIs of 1000, 666.7, 500 and 400 ms respectively). Kuhn found that, although there was no significant difference between the numbers of correct responses for tempo increase and decrease, tempo *decreases* were detected significantly sooner than increases on the reaction-time dependent variable. Surprisingly, there was no effect owing to initial tempo.

In Madsen's (1979) study, subjects were played metronome pulses at initial tempi ranging between 40 and 200 bpm (IOIs of 1500 to 300 ms). These initial tempi increased by 10 bpm over 10 seconds, decreased by the same amount, or stayed constant. Musician and non-musician subjects had to identify how the tempo changed, and were also asked to estimate the final absolute tempo in metronome markings. Both musicians and non-musicians identified tempo decreases significantly more accurately than tempo increases. On the second presentation of each trial condition musicians responded more accurately than non-musicians, suggesting they could learn tempo discrimination more quickly. All subjects tended to underestimate the absolute final tempo, although musicians were generally more accurate.

Wang (1984) had college music students listen to different synthesized versions of the first 16 bars of Dvořák's "Humoreske," Op. 101, No.7.[5] Factors in this study were tempo direction (increase versus

---

[5] An earlier, oft-cited study by Wang (Wang 1983) is, on closer inspection, inconclusive and probably invalid. In the 1983 study Wang played computer-generated versions of four standard-repertoire keyboard

decrease), rhythm (dotted [] versus even []), location of tempo change (eighth-note beats 1, 2, 3, or 4), and texture (accompanied versus unaccompanied). The initial tempo in each case was ♪ = 72 bpm and tempo changed by +/- 1 bpm per bar until the end of the excerpt. Wang found that tempo change was more quickly detected with *decrease* rather than *increase*, with an *even* rather than a *dotted* rhythm, and with an accompaniment rather than without. In addition, although location was not significant as a main effect, it formed a significant interaction with each of the other three factors: whereas *decrease* was detected sooner than *increase*, and *even* sooner than *uneven* on beats 1, 2 and 3, on beat 4 the *tempo-increase-even* rhythm was detected sooner than the other conditions. This highly conditional result is hard to explain in detail, but it does point to the possible influence of metrical location on the detection of temporal change. Finally with regard to Wang's 1984 experiment, it is also interesting to note that tempo change detection was easier when a succession of *isochronous* IOIs was presented, either as an even melody rhythm or as regular, repetitive accompaniment figure.

In two related studies, Geringer and Madsen (1984) and Madsen, Duke and Geringer (1984) examined (among other factors) the detection of tempo difference between pairs of extracts of recorded music. Geringer and Madsen (1984) played orchestral excerpts to a mixture of musician

---

pieces, and asked subjects to detect different types of tempo change that had been incorporated into each rendition. Unfortunately, Wang did not control for starting tempo, rate of tempo change, or rhythm and subdivision of the beat (activity) across the different pieces. Whilst Wang's (1983) use of real compositions is laudable, the untrammelled between-piece differences make her results uninterpretable.

and non-musician subjects, whilst Madsen et al (1984) played extracts of wind-band music to wind and percussion players. In both studies initial tempo was that given on the recording used, and therefore this source of variance was not controlled. However, tempo difference between pairs was always +/- 12%. Both experiments found that subjects identified the type of change significantly more accurately when the second excerpt of a pair was *faster* rather than slower than the first excerpt. In other words, in contrast to Kuhn (1974), Madsen (1979), and Wang (1984), tempo increases were detected more easily than decreases of the same magnitude.

A more recent experiment indirectly provides evidence that tempo decreases are more detectable. Duke (1994) played non-musicians pairs of short rhythms at contrasting tempi; subjects had to say whether the rhythms were the same or not. A tempo of 100 bpm (600 ms) was compared with a contrasting tempo of either 50 bpm (1200 ms), 75 bpm (800 ms), 125 bpm (480 ms) or 150 bpm (400 ms). The 100-bpm tempo could occur before the contrasting tempo, or vice versa. Subjects were better able to equate rhythms when the tempo of the second rhythm was faster than the first, suggesting that tempo increases are less noticeable than decreases.

Generally, whereas some of the above studies suggest that tempo decreases are more easily detected than increases, other studies suggest the opposite. Similarly contradictory results are found among those experiments that have examined the effect of initial tempo on tempo change perception: as noted earlier, Kuhn (1974) found no significant effect owing to initial tempo when using tempi of 60, 90, 120 and 150 bpm. Dornout (1980), however, found tempo change detection easier at 60 bpm than 120 bpm for fourth- through twelfth-graders listening to metronomic pulses. Similarly Wang and Salzberg (1984) found that string students (7 to 18 years old) identified change more accurately when excerpts of string pieces started at slow (60 and 70 bpm) rather than fast tempi. In contrast Wapnick (1980) found tempo change detection for recorded piano music to be more accurately estimated the faster the starting tempo. In the same vein as Wapnick (1980), Ellis (1992) found that third- through sixth-graders' detection of gradual tempo change in a simple synthesized melody was superior at 90 and 135 bpm than at 60 bpm.

New light is thrown on the related problems of change direction and initial tempo in studies by Yarbrough (1987), Ellis (1991) and Brittin (1992; 1993). Yarbrough (1987) played musician and non-musician subjects four recordings of solo piano music by Mozart and Chopin (a "slow" and a "fast" movement by each composer). Pieces were presented in pairs with either no change in tempo across a pair or with a tempo change of +/-18%. Subjects most accurately identified tempo decreases for the slow pieces and tempo increases for the fast pieces.

In a more controlled experiment, Ellis (1991) sought to determine the exact thresholds for tempo change detection across a range of starting tempi. Ellis played subjects a simple monophonic melody that either accelerated or decelerated in tempo, and that began at one of six starting tempi between 48 and 228 bpm (1250 to 263.2 ms). Ellis found that at slow to medium tempi (48 and 84 bpm) decelerations were detected more easily than increases, at medium to fast tempi (120 and

156 bpm) increases and decreases were detected with approximately equal ease, and at fast tempi (192 and 228 bpm) increases were detected marginally more easily than decreases. A similar interaction between change direction and initial tempo has been noted in studies on the perception of tempo change of a conductor's beat using purely visual (Brittin 1992) as well as audio-visual (Brittin 1993) stimuli.

The results of Yarbrough (1987), Ellis (1991) and Brittin (1992, 1993) suggest that subjects are better able to detect decelerations at slow speeds and accelerations at fast speeds. In other words, subjects notice tempo change more easily when tempo modulates away from the central or preferred tempo range identified in the previous subsection of this review. However, although this explanation may well have some merit, it does not match the complexity of the data garnered in the other experiments reported. For instance, Madsen (1979) sampled the entire tempo range from 40 to 200 bpm and found a consistently more accurate response for tempo decelerations regardless of initial tempo. All in all, the complexity of the results cited so far in this subsection suggests that a number of factors come into play in determining the optimum direction and range for tempo change perception. Two factors that may contribute to the incongruity of the results are as follows: (1) the contrasting ways experimenters generated so-called "equivalent" rates of tempo modulation, e.g., +/- 1 bpm per second (Madsen 1979; Wang 1984) versus +/- 12% of the initial tempo (Geringer and Madsen 1984; Madsen, Duke and Geringer 1984); (2) the contrasting ways experimenters have presented tempo changes, i.e., continuous gradual tempo modulation (Kuhn 1974; Madsen 1979; Wang 1984) versus comparison pairs (Geringer and Madsen 1984; Yarbrough 1987). A third stimulus-related confound that will be discussed briefly in the following paragraph is the varying complexity of the stimuli presented in the different experiments.

The literature reviewed may be divided into experiments that use metronomic or computer-generated stimuli, and those that use recordings of real music. Kuhn (1974), Madsen (1979), Wang (1984) and Duke (1994) use simple metronomic or computer generated stimuli, and all these studies reported easier detection for decelerations. In contrast, Geringer and Madsen (1984) and Geringer, Duke and Madsen (1984) played subjects manipulated recordings of real-life orchestral and band music, and they reported easier detection for accelerations. Evidently, the more complex (and musically realistic) a stimulus is, the less controlled are such factors as beat subdivision, varying rhythmic activity, melodic and harmonic activity, and local rubato and agogic changes of note duration.

In fact, a small number of experiments investigating the effect of melodic, harmonic, rhythmic and metric factors on tempo change perception have been carried out. In addition to Wang's (1984) paper reported above, two further studies by Kuhn (1987) and Duke (1989a) explored the effect of melodic/rhythmic activity on the identification of tempo change type between pairs of short synthesized extracts. Kuhn (1987) manipulated three variables: tempo (92 bpm [652.2 ms] or 108 bpm [555.6 ms]), meter (quadruple or triple) and ornamentation (either a melody was played with a simple isochronous quarter-note rhythm, or embellishments consisting of eighth and sixteenth-notes were added).

Kuhn found no effect owing to meter.[6] However, presence of melodic/rhythmic ornamentation *did* influence perception of speed, regardless of the differences between beat rate. This second finding was essentially replicated by Duke (1989a) who played subjects paired variations of a theme by Holst at two tempi (100 bpm [600 ms] and 112 bpm [535.7 ms]). Again, tempo judgments seemed to depend on melodic/rhythmic activity.

Kuhn's (1987) and Duke's (1989a) findings apparently question the stipulation given in Chapter 1 of this dissertation that tempo is not to be equated with rhythmic activity. However, some qualification of Kuhn's and Duke's results is in order: subjects in both studies were mostly second- through sixth-graders (Duke used some undergraduate nonmusic majors), and, even though Duke at least took pains to introduce participants to the abstract notion of beat, it is still not clear whether subjects were aware of the distinction between beat tempo and rhythmic activity on the conceptual level. Both experimenters interpret their results from the point of view of music educators, and doubtless their findings have some pedagogical value (beat is a more abstract concept than melodic/rhythmic activity). However, to show convincingly that tempo is consistently confused with rhythmic activity, one would need to obtain similar results with professional musicians.

The general findings of this subsection on tempo modulation perception may be briefly summarized. The majority of experiments using metronomic or other carefully controlled stimuli suggests that decreases

---

[6] Kuhn's (1987) negative result regarding meter is supported by the findings of a pilot study by the present author reported in Appendix A.

(decelerations) in tempo are easier to detect than increases across a broad range of initial beat rates. Exceptions to this generalization fall into two general categories: (1) two studies using recordings of real music, with all the uncontrolled timing variables that implies (Geringer and Madsen 1984; Madsen, Duke and Geringer 1984); (2) a group of studies that suggest that decelerations are easier to detect at slow tempi and accelerations are easier at fast tempi. Meter seems to have no effect on tempo modulation perception.

Duration, Beat and Meter

The literature that comes to bear on rhythm modulation perception is more disparate than that on tempo modulation. The reason for this disparity is that, whilst no work of which the present author is aware concerns itself directly with rhythmically modulating stimuli, a number of areas of research in psychophysics and music psychology are generally germane to the detection of rhythm modulation. This section is therefore divided into three subsections: (1) a review of some of the empirical work on *just noticeable differences* (JNDs) for discrimination of short durations in audition; (2) a review of the literature concerned with duration cognition using beat-based and metrical models; and (3) a survey of literature concerned with categorical perception and performance of rhythm.

Duration Discrimination and Weber's Law

One oft-cited model of stimulus discrimination in psychophysics that has found currency in duration perception is Weber's law. Weber's law simply states that the just-noticeable difference (JND) for a duration *d* will be a constant fraction of *d* regardless of *d*'s absolute value. Thus if the JND for duration discrimination is found to be, say, 6%, the JND for a sound 1000 ms long would be 60 ms, whereas the JND for a sound 400 ms long would be 24 ms.

In general, psychophysical experimentation on duration discrimination in audition suggests that Weber's law (or a modified version of it [Getty 1975]) holds for moderate durations but not for extremes.[7] Blakely (1933; cited in Woodrow 1951) found that for empty intervals the Weber fraction (JND/base duration) varied only between .08 and .1 for base durations of 200 to 1500 ms, but that the JND rose sharply to as much as .2 below and above this base-duration range. Similar results were found by Stott (1933; cited in Woodrow 1951) for filled durations. Getty (1975) used pairs of empty auditory time intervals ranging from 50 to 3200 ms and found that a Weber fraction of between .05 and .1 accounted for most of the duration discrimination up to 2000 ms. Above this value, however, the JND increased out of proportion to the standard duration. Likewise, Halpern and Darwin (1982) found a mean Weber fraction of .054 for discrimination of durations ranging between 400 and 1450 ms, when the durations were set in a short isochronous rhythmic context. In addition, a number of recent studies that have concentrated on short durations suggest that Weber's law breaks down in the range below about 200 ms: Schulze

---

[7] It should be noted, however, that other data argue against Weber's law, or at least restrict its generality. For example, Allan and Kristofferson (1974) found that the absolute JND remained constant over a wide range of base durations for well-practised subjects.

(1989b), for example, found a larger differential threshold (in ms) for a standard of 100 ms than for 200 ms, and Hirsh, Monahan, Grant and Singh (1990) found increasing Weber fractions for durations under 200 ms.

It is perhaps no coincidence that the duration range in which Weber's law seems to apply, and in which the Weber fraction is lowest, is similar to the range of preferred beat tempo: the above experiments report a fairly constant Weber fraction within the range of 200 to 2000 ms (expressed as metronome markings as 300 bpm to 30 bpm). This correspondence is not surprising, for if the beat level is to form the kernel of temporal structure in a piece of music, the time-interval level around which all else is structured, one would expect that time-interval level to lie in a range where the human auditory system is particularly sensitive to temporal change. In this way, any intended nuances in beat length (agogics, rubato, tempo changes), as well as accidental inconsistencies are very likely to be perceived.

At this juncture, however, it is as well to stress the distinction between a raw perception and a more abstract permanent representation in cognition. On the one hand, as some of the previously reported experiments suggest, simple perception of time intervals is fairly fine-honed (see for example Ellis's 1991 study on tempo discrimination thresholds, and the above literature on Weber's law). On the other hand when it comes to a more reliable, long-term representation of time intervals (reliable enough for reproduction at a later time, for example) another body of studies suggest that the human auditory system is surprisingly crude.

Beat and Meter in Time-Interval Cognition

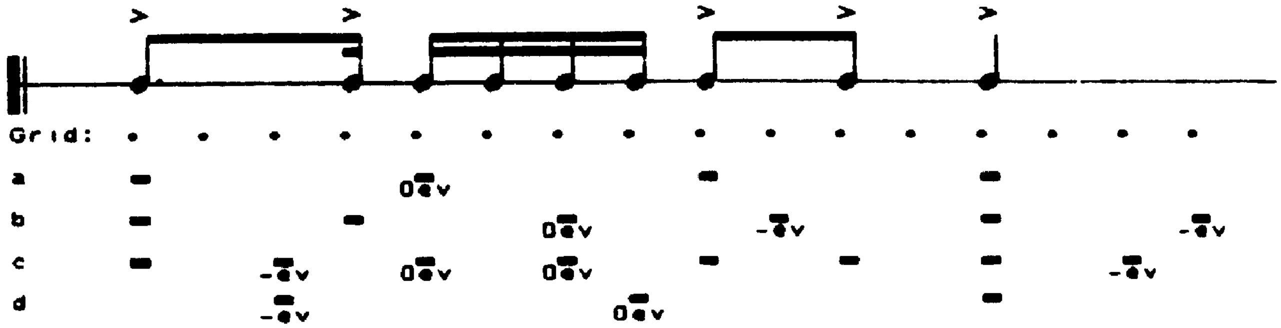
Some studies suggest that listeners are unable to discriminate more than a very small number of time intervals. Fraisse found that in simple tapping tasks, subjects tended to reproduce two types of time intervals, what he called *temps court* (200 to 300 ms) and *temps long* (450 to 900 ms) (Fraisse 1982, 167). Povel (1981), carried out two experimental studies in which subjects had to synchronize with or reproduce (tap back) two- or three-element repeating rhythmic units. The IOI ratios in these experiments were 1:4, 1:3, 2:5, 1:2, 3:5, 2:3, 3:4 and 4:5. Only the 1:2 ratio was reproduced accurately, with reproductions of the other ratios drifting towards 1:2. Povel however reasoned that these findings do not adequately explain how musicians frequently perform much more complicated ratios and therefore are presumably able to discriminate among more than two simply differentiated time intervals. Povel's third (1981) experiment therefore placed time-interval ratios such as 1:3, 1:4 and 2:3 within a more complex context. Subjects' responses were significantly more accurate in contexts where a regular beat could accommodate both the short and long time intervals. Thus, for example, reproduction by musically trained subjects was accurate for the rhythm 250-250-250-750 ms (with a beat of 750 ms) but poor for the rhythm 200-200-600.

To explain the often fine-grained production or reproduction of rhythmic time intervals by performing musicians, Povel (1981; 1984) proposed a beat-based model of rhythm perception. From the definition of beat offered in Chapter 1, beat is an isochronous series of time spans around which are organized all intervals in the stimulus. Thus, in a beat-based model, all stimulus time intervals are "rationalized" as either integer multiples or divisors of the beat time interval. In addition to a specified duration, the beat is also determined with respect to its position, that is, its phase relative to the onsets of events in the stimulus rhythm.

Povel and Essens (1985) propose a model of beat placement in which the combination of beat length and position that is finally chosen is that which coincides with the greatest number of events and accents in the stimulus. Since the stimuli envisaged by Povel and Essens in this model are unpitched and undifferentiated as to loudness level, "accent" is taken as arising from purely temporal cues. The authors take the results of an earlier experiment (Povel and Okkerman 1981) as the basis for identifying accent: accents are perceived (1) on temporally isolated tones, (2) on the second tone of a cluster of two tones, and (3) on the initial and final tones of a cluster of three of more tones. To determine whether tones are isolated or clustered, the onset points of all tones are fitted onto a temporal grid whose divisional units are the shortest time interval inherent in the stimulus. Once accented and unaccented stimulus events have been determined relative to the temporal grid, every possible beat length/position is evaluated for plausibility according to the following formula: $C = (W*-ev) + (1*0ev)$, where C is the counterevidence against a particular beat length/position, -ev is the number of hypothesized beats coinciding with silences, 0ev is the number of hypothesized beats coinciding with unaccented attacks, and W is a constant (set by Povel and Essens to 4).

Figure 2.2 shows four possible beat length/positions for a simple rhythmic stimulus. Although Povel and Essens' model finds the least counterevidence for the beat length/position that is intuitively most plausible (beat "a," a quarter note on the first, fifth, ninth etc. grid units), a couple of criticisms are in order. Firstly, because Povel and Essens' formula only takes into account countervaling evidence, that is beats that fall on a silence or an unaccented attack in the stimulus, the fewer the number of hypothesized beats for a given rhythm, the lower will be the value of C. In other words, the longer the beat length, the less likely a beat will coincide with -evs or 0evs. The undue preference for long beats results in the intuitively implausible beats "b" or "d" being preferred over the more plausible beat "c" in Figure 2.2.

A second criticism of Povel and Essens' model is more general. Whilst their model might provide an interesting and fruitful approach for computerized beat assigment, it is less realistic as a perceptual model. The goal of the model, finding a beat that best corresponds to accents in the stimulus, seems eminently sensible; it is merely the means to that end that are less feasible. To start with, an entire rhythm would have to be stored in some type of sensory buffer before the listener could begin to decipher the beat. In addition, every beat position/length combination would then [illegible] to be tested to see which was most plausible. This computationally demanding *modus operandi* runs counter to the ease with which musical listeners are able to pick up a



$C = (4 * -ev) + (1 * 0ev)$

	-ev	0ev	C
a	0	1	1
b	2	1	9
c	2	2	10
d	1	1	5

Figure 2.2. Four different beat lengths/positions applied to a simple rhythm (after Povel and Essens [1985]). Accents are assigned according to Povel and Okkerman (1981); the grid represents the highest-common-factor duration in the stimulus (a sixteenth note). Beat a is four grid units long (a quarter note), beat b is three grid units long (a dotted eighth note), beat c is two grid units long (an eighth note) and all are positioned in phase with the first note of the rhythm. Beat d is five units long and positioned out of phase with the first note of the rhythm. The table underneath shows that beat a has the least countervaling evidence (the lowest value of C) and therefore is considered the most plausible beat length/position by the model.

beat (e.g., tap their toe) after maybe only 3 or 4 notes of a rhythm have sounded, often without much conscious effort.[8]

Less unrealistic demands are made of the perceptual system in a series of models of meter perception proposed by Longuet-Higgins and Lee (1982; 1984) and by Lee (1985; 1991). Since these studies model meter rather than beat perception, they assume responsibility for assigning not only the "correct" beat length/position, but also attempt to posit intuitively plausible subdivisions and groupings of that beat too. Longuet-Higgins and Lee's (1982) model has as its principal aim the "correct" positioning of the barline. Using purely temporal information, this model takes the first IOI in a stimulus as a hypothetical beat length/position. The model then modifies this beat length/position in the light of subsequent time intervals. The length of the hypothesized beat is doubled until the beat is no longer consistently multipliable. The position of the beat is revised so that *longer* notes are arranged so as to occur at the beginning of the bar or at other metrically important points. The combination of these two types of processes (length and position) result in an unequivocal judgment as to the most likely position and length of a bar for any rhythmic pattern.

In their 1984 paper, Longuet-Higgins and Lee take a slightly different tack. Here the principal rule in beat assignment is that the beat length/position chosen entails as little syncopation as possible.

---

[8] Note, however, that Smith (1983) found that ability to assign beat to rhythmic stimuli was influenced by musical training: musically untrained subjects organized rhythms according to simple grouping or counting strategies rather than by beat.

Syncopation is defined in this case as the situation where a note N is followed by either a rest or a tied note, and where the rest or tied note is at the same or at a higher metrical level than N. As in Longuet-Higgins and Lee (1982), the avoidance of syncopation in most cases means the avoidance of long notes at metrically weak points. Lee (1985) is essentially a modification of Longuet-Higgins and Lee's (1982) model, allowing that model to assign more than one metrical level but disallowing any metrical level longer than the longest note in the rhythmic stimulus. Lee's (1991) model synthesizes some of the preceding models mentioned above. In general the model begins by assuming that the very first sounded note is a downbeat (although this assumption can be revised later), it assigns long notes to downbeats (or other metrically important points), and can infer metrical grouping of the beat as well as beat subdivision. In addition, an important advance in Lee (1991) on previous models is that it takes absolute tempo into account, such that a beat ("tactus") with a time interval in the range of 300 to 600 ms (200 to 100 bpm) is preferred. This beat-rate range is a little faster than that found earlier in this chapter from a review of the literature on preferred tempo, however it is apparently supported by the data of three perceptual experiments performed by Lee (1991).

As valuable as the above models are, one often-acknowledged limitation with them is that they concentrate almost exclusively on temporal cues for beat and meter. A growing number of studies support the common-sense notion that various other types of accents, what Lerdahl and Jackendoff (1983, 17) call "phenomenal accents," can also act as cues to metrical structure. In order to establish the length/

position of higher levels of the metrical hierarchy, the listener makes use of the placement of the following accent types: dynamic stress (Drake, Dowling and Palmer 1991; Windsor 1993), melodic accent (i.e., a change of contour a leap in a melodic line [Thomassen 1982; Drake et al 1991]), or change of underlying or implied harmony (Dawe, Platt and Racine 1993; Smith and Cuddy 1989).

Parncutt (1994) has developed a quantitative model of beat and meter assignment that, in principal at least, can take various types of phenomenal accents into account. In practice, however, Parncutt has so far limited the testing of his model to durational accent alone. In this model different isochronous templates are matched to a complete input rhythm. The template that (a) matches the greatest number of events and (b) best coincides with a preferred-time-interval range centred on 700 ms is chosen as the most likely beat length/position. To assign a downbeat, the model chooses the position and integer multiple of the beat length that most coincides with the placement of phenomenal accents in the rhythm. Parncutt's model is reminiscent of Lee's (1991) model in that it takes tempo range into account, but is more like Povel and Essens' (1985) model in that it handles complete rhythms rather than working from left to right "in time."

All of the models so far considered, whether they work on complete rhythms (e.g., Povel and Essens 1985) or from left to right (e.g., Lee 1991), tacitly assume exact, integrally related time intervals as their input. The following subsection explores the way in which the human auditory system is able to handle the types of inexact, non-integrally related time intervals that abound in musical performance.

Categorical Perception

Categorical perception is the process whereby the infinite, continuous stimuli from the physical world are pigeon-holed into a small, manageable number of discrete categories. For example, the frequency ratio between two tones may be one of an infinite number of different frequency ratios.. However, studies reveal that Western listeners tend to match the perceived relationship between two pitches to one of a limited repertoire of culturally normative musical intervals, i.e., the set of chromatic intervals (Burns and Ward 1978; Siegel and Siegel 1977).

A further example of categorical rhythm perception is provided by comparing two of the bodies of the literature reviewed so far. Experimental work on simple duration discrimination (e.g., Weber's law experiments) reveals a fine-grained, continuous perception of durations. The Weber fraction may take any value and is limited purely by the acuity of the listener's perceptual system. In contrast, those experiments that required the subject to perform a task where time intervals would have to be stored or internally represented for later processing tell a different story. Recall that Povel (1981) found that subjects tended to reproduce non-integer-related rhythms as integer-related. In this case, the infinitely divisible time-interval-ratio continuum was divided up into discrete, integer-related pigeon holes.

In the majority of the experimental studies on categorical perception, two complementary paradigms have been used: identification and discrimination (Handel 1989, ch 9). Identification simply involves matching a stimulus to one of two or more labels; discrimination, entails the subject telling apart two stimuli. If two stimuli, A and B, straddle a presumed categorical boundary, then a significant majority of the subjects' responses (conventionally chosen as 75% or more) will give stimuli A and B different labels on the identification task. Conversely, if A and B fall within one category, then they will identified as equivalent at least 75% of the time. By the same token, discrimination between A and B will be easiest if they straddle a categorical boundary, and less easy if they fall within the same category, regardless of the physical difference between A and B.

To test explicitly for the existence of categorical perception of rhythms in a metrical context, Clarke (1987a) devised two related experiments, one of perception the other of perception and performance. In the first (perceptual) experiment music students heard the short melodies shown in Figure 2.3 (Clarke, 1987a, 23). The first melody established a compound meter (with three subdivisions per beat), the second melody a simple meter (with two subdivisions per beat). The notes marked X and Y were of variable duration and given one of the pairs of IOIs (in ms) outlined at the foot of Figure 2.3.

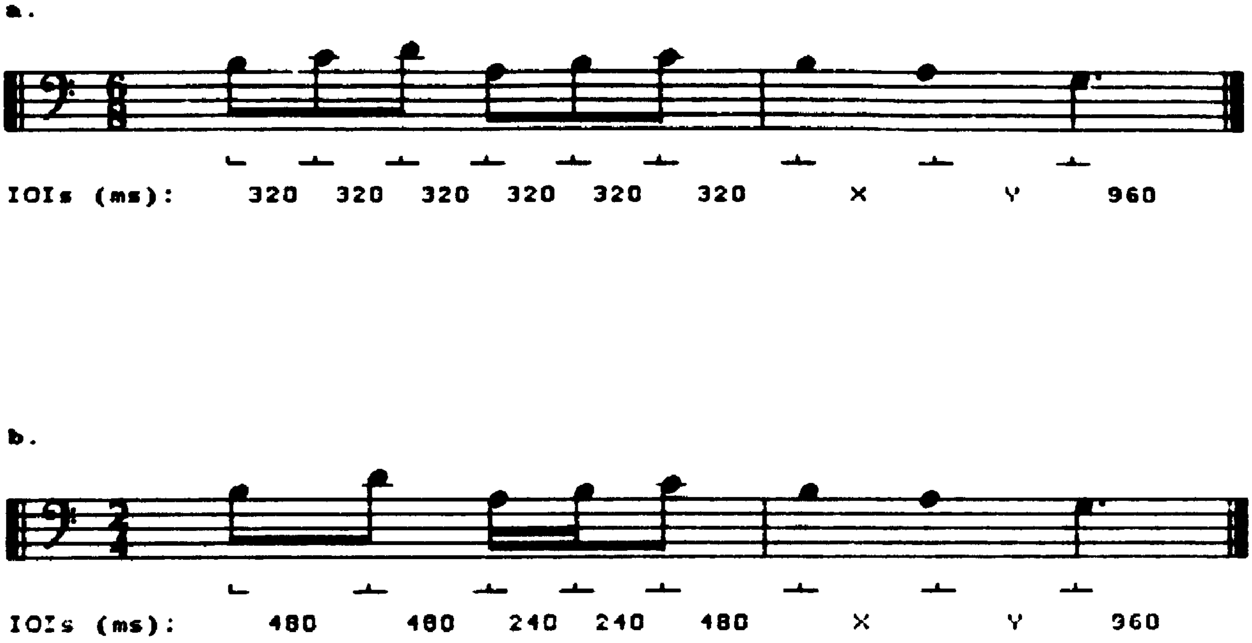
Subjects performed both identification and discrimination tasks. For the identification task subjects had to say whether the notes X and Y matched either the ♩♪ or the ♪♩ rhythm. For the discrimination task subjects heard pairs of sequences separated from each other by two

SEQUENCE	X (ms)	Y (ms)		NOTATION
1	640	320	=	♩ ♪
2	620	340		
3	600	360		
4	580	380		
5	560	400		
6	540	420		
7	520	440		
8	500	460		
9	480	480	=	♪ ♪

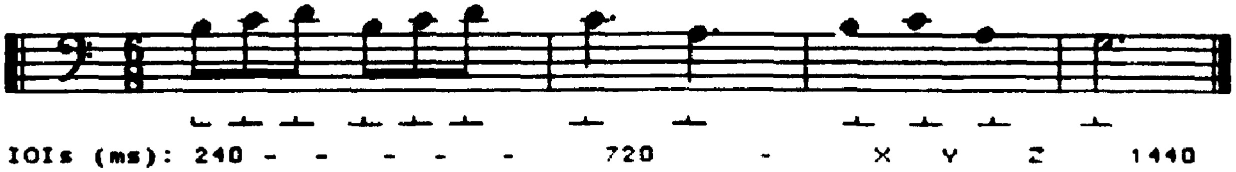
Figure 2.3. Clarke's (1987a) perceptual stimuli. Figure a shows the 6/8 context, Figure b shows the 2/4 context. The values of the variable durations X and Y are given in the table at the foot of the figure. Only sequences 1 and 9 are notatable, with 1 conforming to a 6/8 meter and 9 conforming to a 2/4 meter.

positions on the scale of X-Y durational variants (thus sequence 1 was compared with 3, or 6 with 8, for example).

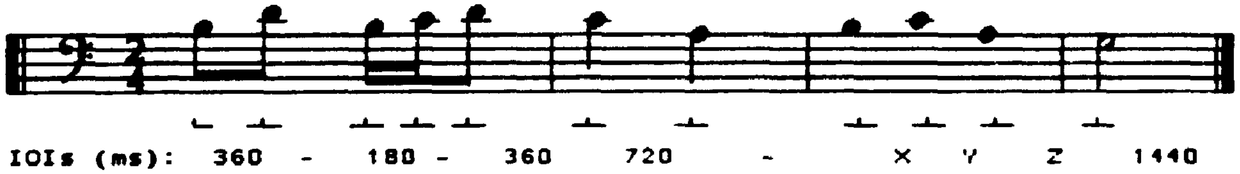
In the identification task, subject responses showed an abrupt shift between the unnotatable sequences 5 and 7, with around 80% of the responses identifying 5 as ♩♪ and around 80% identifying 7 as ♪♪. In addition, in the discrimination task, subjects found the pairing of sequences 5 and 7 easiest to tell apart, indicating that these two stimuli straddle a categorical boundary. Clarke also partitioned the results of his perception experiment to see if the type of metrical context (compound or simple) affected the position of the categorical boundary. Generally, in the compound metrical context a greater proportion of the X-Y time-interval values were categorized as ♩♪. Conversely, in the simple metrical context, a greater proportion of the X-Y time-interval values were categorized as ♪♪. In short, ambiguous rhythmic values were categorized as conforming to the prevailing metrical context.

As a complement to the perceptual study, Clarke (1987a) also carried out a performance experiment where pianists were played one of the melodies shown in Figure 2.4, and then were asked to play the melodies back on a piano connected to a computer. The IOIs of the notes in each stimulus marked X, Y and Z were variable according to the values given in Figure 2.4.

From the variance of the performance timing data Clarke concluded that: (a) the most stable (reliable) and accurate rhythms tended to have low-value integer time-interval ratios; (b) the least stable and accurate rhythms were unnotatable and had complex time-a.



b.



SEQUENCE	IOIs (ms)			PROPORTIONS
	X	Y	Z	
1	480	480	480	1:1:1
2	420	420	600	
3	360	360	720	1:1:2
4	420	300	720	
5	480	240	720	2:1:3
6	510	210	720	
7	540	180	720	3:1:4

Figure 2.4. Clarke's (1987a) performance stimuli. Figure a shows the 6/8 context; figure b shows the 2/4 context. The table at the foot of the figure gives the IOIs of notes X, Y and Z as well as the IOI ratios when those proportions are integrally related.

interval ratios; (c) an exception to (b) was sequence 6. Taking sequences 5, 6 and 7 together, the X and Y durations of these three rhythms may be regarded as variants of a basic "long-short" rhythm. Sequence 5 conforms to the 6/8 metrical context, and is indeed more stable and accurate in that context. On the other hand, sequence 7 conforms to the 2/4 metrical context and is reproduced more stably and accurately in that context. Rhythm 6 represents a mid-way point between rhythms 5 and 7 and is reproduced moderately stably and accurately in both the 6/8 and the 2/4 contexts. As was the case in Clarke's first (1987a) experiment, these results suggest that metrical context influences the category to which ambiguous rhythmic stimuli are assigned.

To expand on Clarke's paradigm, Schulze (1989a) had subjects learn the four archetypal rhythms shown in Figure 2.5 and a number of variants that sampled durational continua between and around these rhythms. Subjects had to identify rhythms simply by matching the stimulus to a learned numerical label. In addition, to see if categorical perception held for more than one tempo, Schulze randomly varied the stimulus pattern-repetition rate across trials. Generally the archetypal, integrally related rhythms were identified more accurately than non-integrally related variants. In addition, discrimination evinced a sharp peak between variant rhythms about half-way along each continuum, suggesting a categorical boundary at that point. Again, these results point to categorical perception of simple repeated rhythms.

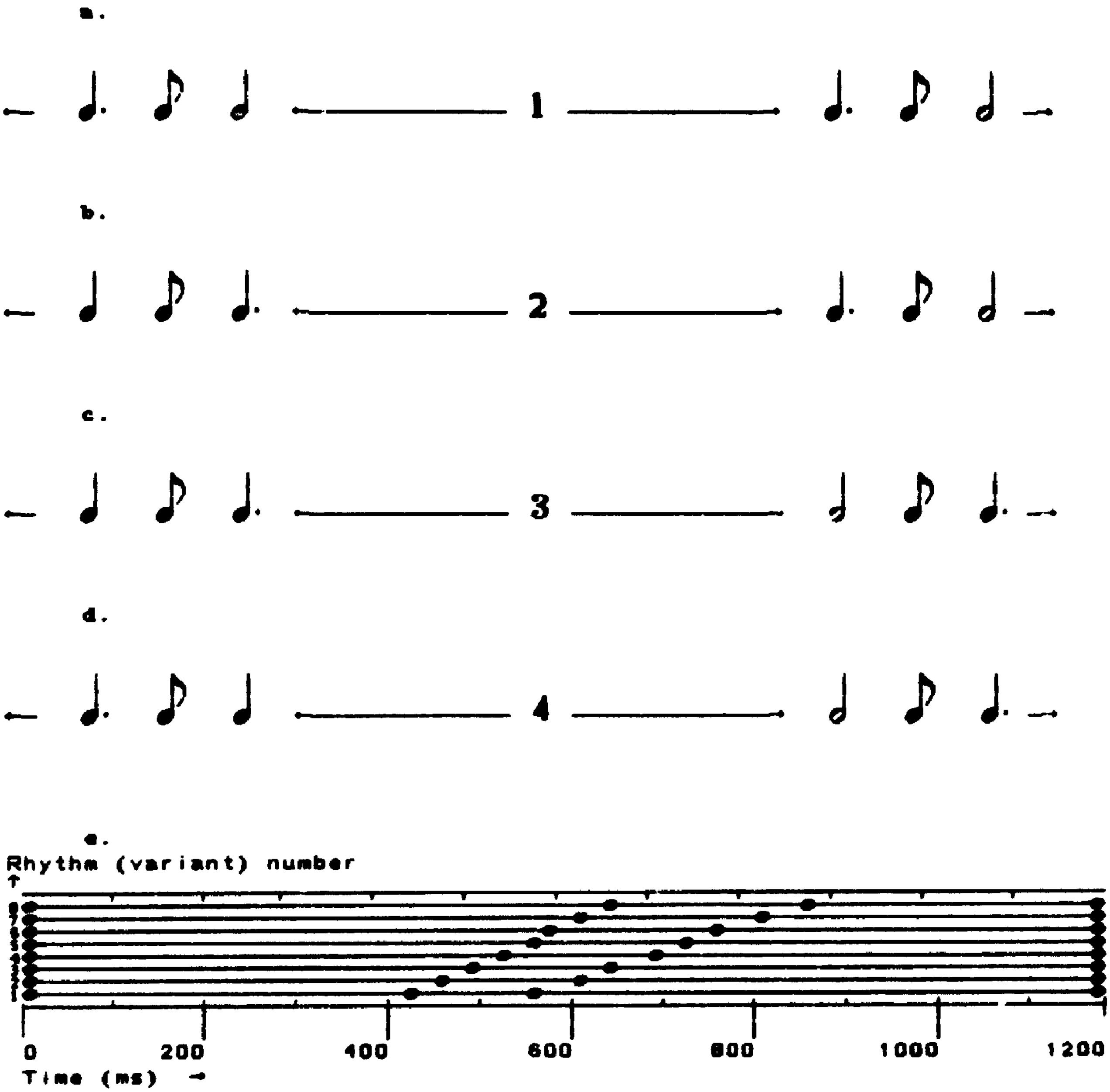


Figure 2.5. Schulze's (1989a) stimuli. Figures a through d show the rhythmic continum (1-4) established between and around four archetypical integrally related rhythms. Figure e shows how the rhythmic continuum number 1 (figure a) was sampled to give eight rhythms between and around the archetypical rhythms: time (IOIs) in ms is given on the x-axis, the number of the variant rhythm is given on the y-axis, onsets of notes are shown by the • symbol. In this case, the duration of one pattern is 1200 ms, regardless of the IOIs between individual onsets. Rhythms 2 and 7 are the archotypical rhythms.

Given that beat rate was randomly varied, Schulze also took his results to be a generalization of Clarke's findings across the tempo range. Unfortunately, however, Schulze used the following three pattern rates: 1200 ms/pattern ("medium" as exemplified in Figure 2.5e), 1320 ms/pattern ("slow") and 1080 ms/pattern ("fast"). Taking each pattern to be two beats long, these rates translate into the following tempi: medium = 60000/600 = 100 bpm; slow = 60000/660 = 91 bpm; fast = 60000/540 = 111 bpm. Clearly a wider sampling of the musical tempo range is required if the generalizability of categorical rhythmic perception across the musical tempo range is to be empirically substantiated.

The experiments considered so far have used computer-generated stimuli that present IOIs with precisely delineated ratios. On the other hand, as noted in Chapter 1, in human musical performance such precision is practically impossible; the time-interval ratios between adjacent notes rarely match their "notional" values. If human-generated rhythms are almost never precisely executed, how does one explain the above-mentioned preference for integer-related time-interval ratios and categorical perception?

To answer the question posed above, first it would be best to see which systematic modulations of exact metronomic rhythm normally occur in expressive, well-performed music. A number of researchers have noted how structurally important events (phrase boundaries for example) are lengthened in performance (Gabrielsson 1974; Povel 1977; Seashore 1967; Sloboda 1985). In addition, Shaffer (1981) found that in a performance of Bach's *alla breve* fugue in Eb from Book I of the *Well-*

Tempered Clavier, the pianist systematically lengthened the IOIs of the first and third quarter notes in each bar. Along similar lines, Sloboda (1983) observed that notes located at the beginning and half way through a duple measure are played louder and longer from onset to offset of the tone.

To account for these phenomena, Todd (1985) advanced a model of rhythmic performance stipulating a greater positive deviation (IOI lengthening) the higher a note is in the metrical and hypermetrical hierarchy. In the same vein, Clarke (1985) proposed a model of expressive timing in musical performance in which the isochronous pulsing of a motor control clock is itself modulated by higher-order control curves that reflect the structure of the music to be played. In this way the degree of expressive timing at any point in a performance reflects the structural (hypermetrical) importance of the musical events at that point.

The above models, then, are able to reconcile the non-integrally related IOIs of musical performance with listeners' preference for integrally related time-interval ratios: All assume the existence of some type of clock that pulses at a constant rate. Integrally related rhythms are thus easily accommodated as either integer divisors or multiples of the period of one clock pulse. In this last sense, the models of Todd (1985) and Clarke (1985) are reminiscent of the models of Jones (1976; Jones and Boltz 1989) and Povel (1984) reviewed earlier in this chapter. They differ, however in that this constant beat rate is itself modulated in order to convey the relative structural importance of events within a musical piece.

Clarke's and Todd's models are intended to be generative, that is, they describe musical performance. Clarke (1985) assumes that perception and categorization of expressively modulated stimuli would simply involve the generative process in reverse. Although a further (1989) experimental study by Clarke provides evidence that listeners are sensitive to temporal changes within an expressively modulated musical phrase, Clarke does not attempt to explain how a stable representation of beat and/or metrical structure might be extracted from such a modulated stimulus. An account of this relationship is one of the goals of the model outlined in the next chapter, Chapter 3.

CHAPTER 3

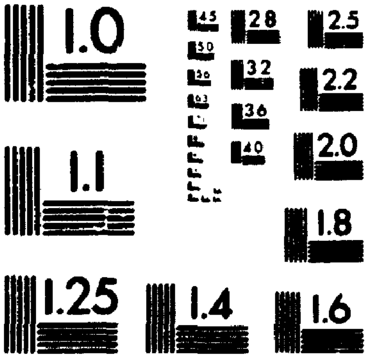
MODEL OF SHORT TIME-INTERVAL PERCEPTION IN MUSIC

This chapter presents a model of the perception and cognition of short time intervals as they might be processed by the human brain in the auditory modality. Although all sounds reaching the tympanic membrane must have duration, the model advanced here is primarily concerned with the types of temporal relationships specific to music. In order to make the model as internally consistent as possible, three guiding principals or goals have been adopted in the construction of the model. These goals are outlined briefly in the first section of this chapter.

Goals of the Model

Firstly, the model attempts to account for as many of the experimental results summarized in the previous chapter as possible. In other words the model seeks to integrate the findings of these experiments into one coherent account.

A second goal of the model is to generate specific, experimentally testable hypotheses. The type of detail required of a model to give rise to particular hypotheses may seem inconsistent with the generality required to address the findings of a wide-ranging literature review. The tug-of-war between generality and specificity is


1.0
4.5
2.8
2.5
5.0
3.2
2.2
5.6
6.3
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

neutralized in the case of this model by focusing on certain perceptual/cognitive stages whilst giving others more cursory treatment. The stages that enjoy special attention are those whose workings are examined experimentally in Chapters 4 and 5 of this dissertation.

The third goal may be approached in an indirect way by considering Agmon's (1990) distinction between *what-* and *how-theories* of music cognition. According to Agmon, what-theories simply stipulate *what* mental behaviours will occur given a particular set of circumstances (for example, a certain stimulus and a certain type of subject). On the other hand, how-theories postulate the means by which an observed mental behaviour arises given a particular set of circumstances. In other words, what-theories describe mental behaviour whereas how-theories attempt an explanation of that behaviour. As Agmon notes, what-theories must logically precede how-theories since one must know what happens in a particular system before one can say how it happens.

Research in rhythm perception is still very much in the what-theory stage. The general thrust of much experimental work in the area concerns itself with what listeners generally hear when presented with particular rhythmic stimuli. Given the limitations of the data so far available, the model presented here necessarily tends towards a what-theory. The model collates findings about *what* the human perceptual system does with particular rhythmic stimuli. Its originality lies in bringing together bodies of research that have, to all intents and purposes, ignored each other so far. However, that having been said, a modicum of how-theorizing cannot but help creep into the discourse of this chapter as well. In the act of integrating observed behaviours, the theorist must also speculate on how particular mental operations might interact. Whilst currently the neural workings behind rhythm perception can only be guessed at, the juxtaposition of cognitive stages in this model putatively suggests causal explanations for some aspects of rhythmic behaviour at a more general level.

The above consideration of what- and how-theories leads to the third goal of the model: to provide a clear conceptualization of what the human perceptual system does with short (rhythmic) time intervals, and to suggest how these operations interact. Many of the terms used (e.g. "oscillator," "filter," "gating" or "logic") are overtly reminiscent of an electronic system. This electronic flavour is not accidental; it is hoped that describing the rhythmic processor in electronic terms will prove both useful and thought-provoking, for by doing so the model's analogical role is highlighted.

Specific Matters to be Addressed

Having framed the principal goals of the proposed model, the discourse now turns to more specific issues arising in the literature review. Work on the perceptual present and echoic memory suggests that temporal information is integrated fairly early in the perceptual system. Research in a number of fields converges on the maximum value for perceptual temporal integration of 2-3 seconds. On the other hand, other evidence suggests that auditory events can be stored for recall and comparison in auditory short-term memory for as much as 20-30 seconds. Any model of time-interval perception should reflect the interaction of these two temporal ranges representing perception and memory respectively.

Secondly, this model will assume that a time-interval processor works by tuning internal (probably neural) oscillators to the periodicities inherent in the incoming stimulus. The hypothetical means by which this entrainment comes about will be described later. It is important to note, however, that such an approach entails the model working in real time, processing time intervals from note to note. In this respect, then, the model presented here is more akin to the models offered by Longuet-Higgins and Lee (1982) and Lee (1991) than those offered by Povel and Essens (1985) or Parncutt (1994).

The complementary relationship between ongoing perception and memory, and the importance of a hierarchy of entrained oscillators determine the general shape of the model. In addition, certain other issues brought to light by experimental work discussed in Chapter 2 must also be addressed:

a. The preferred beat tempo range between roughly 60 and 150 bpm (1000 and 400 ms).
b. Listeners' greater facility for hearing tempo decelerations over accelerations with simple, metronomic or computer-generated stimuli.
c. The interaction between metrical location of tempo change and direction of change as regards the ease of tempo change detection by listeners.
d. The reasonably fine-grained detection of durational anomalies in perception (c. +/- 5%) as found by some of the experiments adopting the Weber's law paradigm.

e. The much less fine-grained reliance on integrally related IOIs in rhythm cognition.

f. The tendency by musically more sophisticated listeners to parse a beat from an input stimulus to provide a reference time interval around which all other time intervals are organized.

g. The rapid establishment of a metrical hierarchy as multiples and divisors of the beat time interval, after only a few notes have sounded.

h. The tendency to categorize time intervals not only as integrally related, but also to categorize them as conforming to an established metrical context.

The Model

A model of time-interval perception is outlined as a block diagram in Figure 3.1. As is evident from Figure 3.1, the input stimulus is immediately split into two channels, implying parallel processing. In this way, two types of time intervals are extracted from the stimulus: (1) simple time intervals between successive individual events registered by the INTER-EVENT INTERVAL TIMER (IEIT), and (2) time intervals between more complex phenomena that elicit a sense of grouping or mark phenomenal boundaries, such as melodic, durational, dynamic or timbral accents, changes of harmony, repetitions of melodic phrases, etc., registered by the INTER-ACCENT INTERVAL TIMER. It is the simpler INTER-EVENT INTERVAL TIMER, the IEIT, and its interaction with the OSCILLATOR BANK that will be examined first.

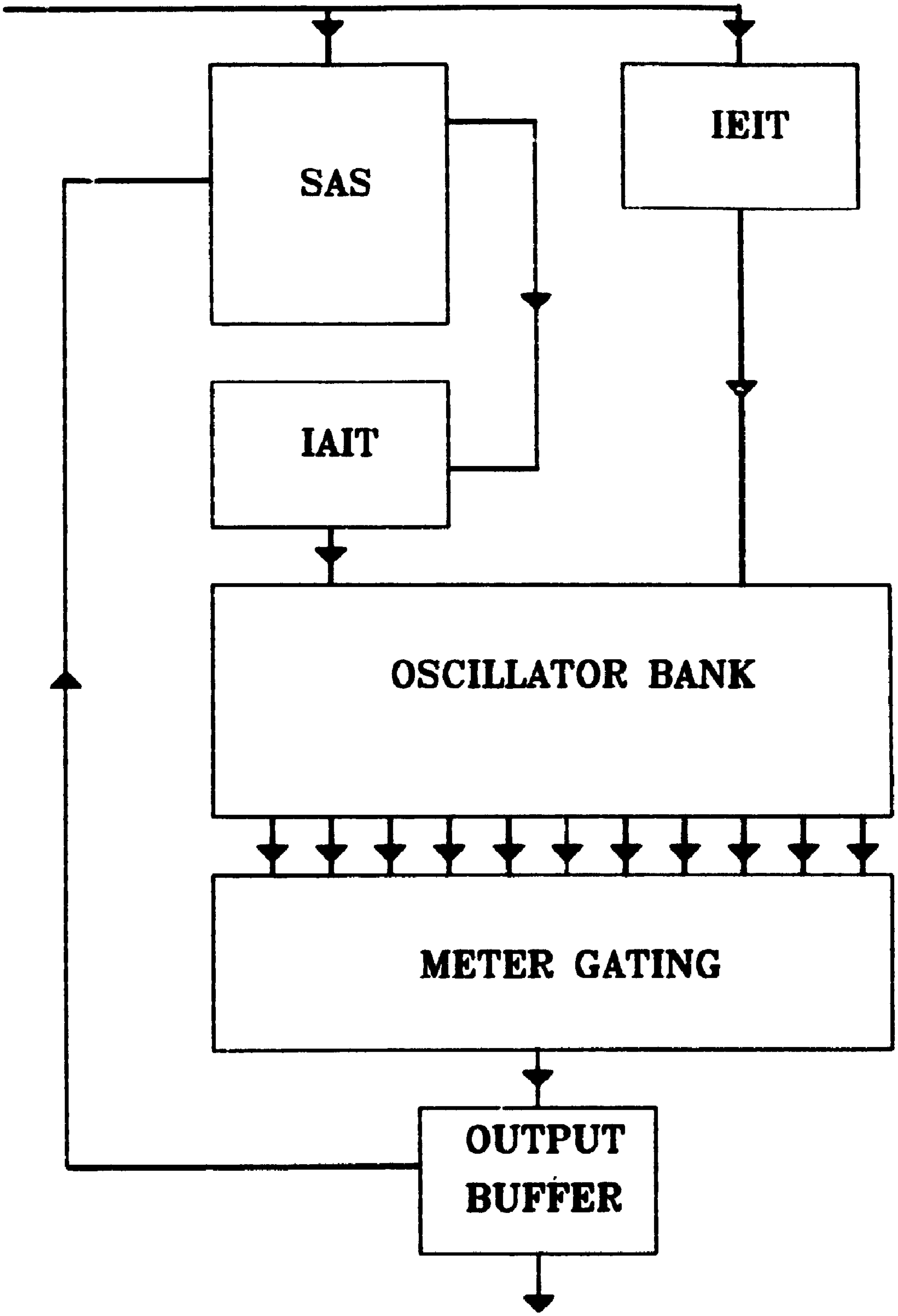


Figure 3.1. Perceptual model for processing short time intervals in music.

Measuring Time Intervals between Events

The IEIT has a simple task: to produce a sharp, spiky energy pulse every time a stimulus event onset occurs. Since, as can be seen in Figure 3.2, the temporal placement of these pulses is clearly defined, the rhythm formed by a string of IEIT pulses may be taken as identical to the stimulus rhythm. A time-span that elapses between stimulus onsets is an IOI. A time-span that elapses between IEIT pulses will be referred to simply as a "time interval." However, IOIs and IEIT pulse time intervals may be taken as synonymous, and the two terms will be freely interchanged in the ensuing discourse.

The OSCILLATOR BANK, illustrated in detail in Figure 3.3, is presumed to contain a small number of neural oscillators that are driven by pulses from the IEIT. These oscillators produce low-frequency, ramp-shape waveforms when their output energy level is measured over time. A typical oscillator output is illustrated in Figure 3.4. The time-span of one cycle, that is, the time-span between two maxima on the waveform shown in Figure 3.4, is referred to as a "period," designated in milliseconds. For the purposes of illustration in this chapter, it is assumed that there are eleven oscillators in the OSCILLATOR BANK with the output period ranges shown in Table I. These ranges, placed on a logarithmic scale, are arbitrarily chosen, however their overall span is derived from the preferred tempo range reliably observed in the literature reviewed in Chapter 2.

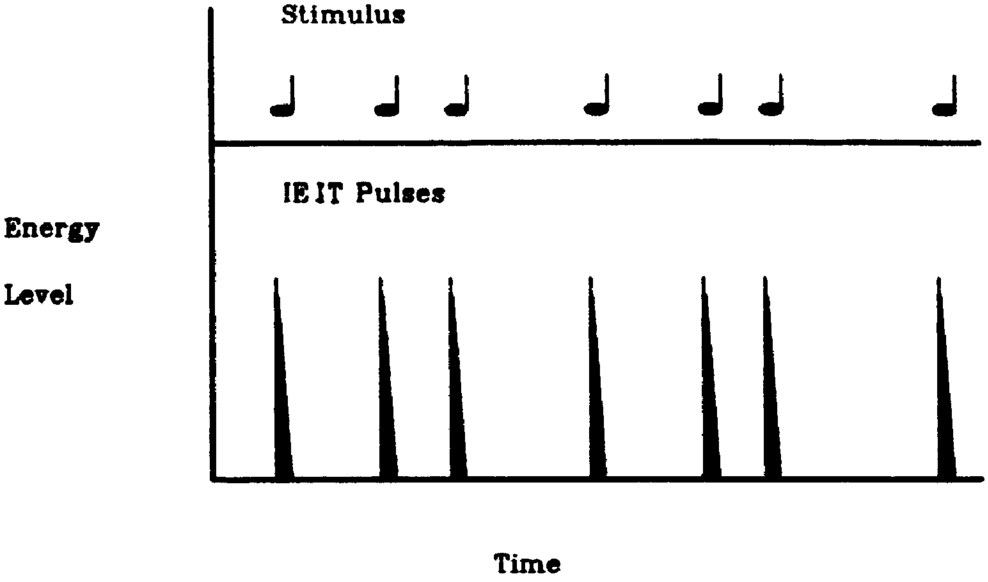


Figure 3.2. Relationship between stimulus onsets and IEIT pulses.

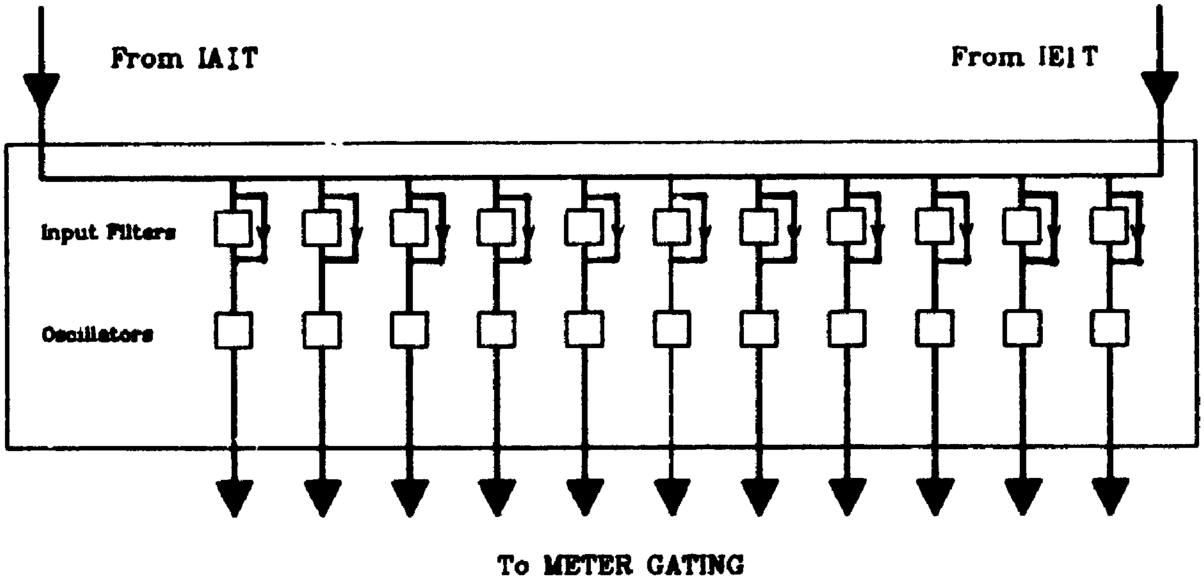


Figure 3.3. OSCILLATOR BANK in detail.

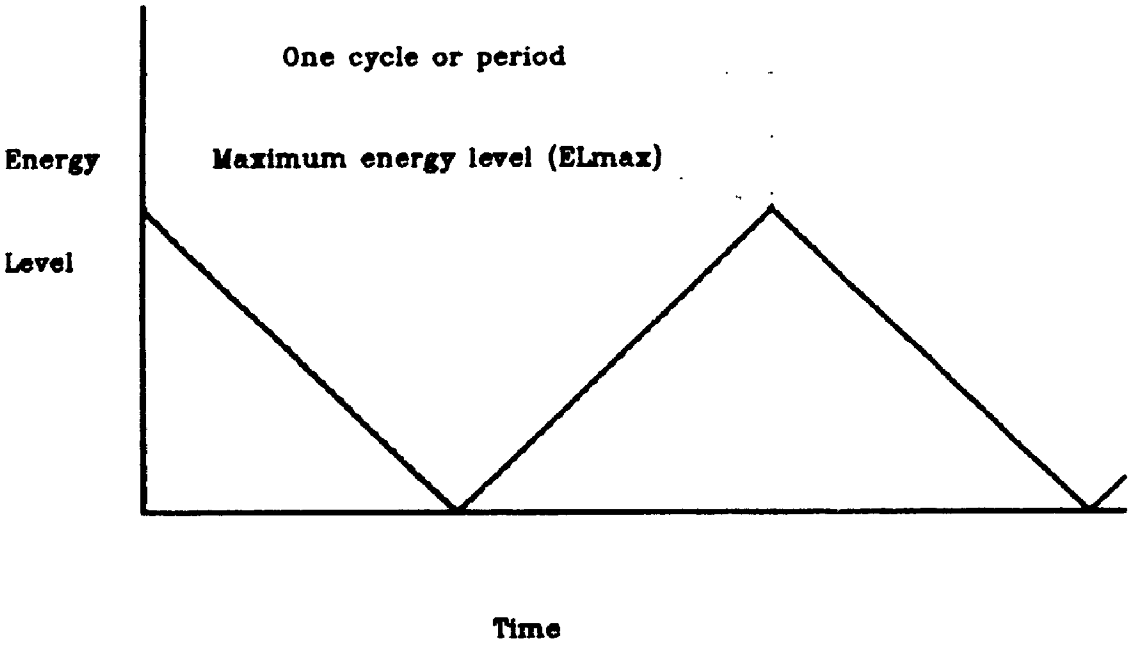


Figure 3.4. General shape of output waveform of oscillator.

TABLE I

OUTPUT-PERIOD RANGES (IN MS) OF OSCILLATORS IN THE OSCILLATOR BANK

Oscillator#	Mean Period	Lower bound	Upper bound
1	200	145	280
2	270	200	370
3	350	260	470
4	430	330	570
5	520	400	680
6	600	470	770
7	690	530	900
8	830	630	1100
9	1040	770	1400
10	1350	980	1860
11	1800	1300	2500

Table I shows the hypothetical upper and lower bounds of each oscillator as well as the mean period (on a logarithmic scale). The ranges given in Table I may also be rendered graphically as a set of bell-curves, as shown in Figure 3.5. From Figure 3.5 it can be seen that, for each oscillator, the mean period is the period that the oscillator produces its highest amplitude. Conversely, the periods at the upper and lower bounds (at +/- 3 standard deviations either side of the mean) can only be produced at minimal amplitudes. However, since the period ranges of the oscillators overlap, a period that can only be produced at a low amplitude by one oscillator will be produced at a higher amplitude by another, neighbouring oscillator.

It should also be noted from Figure 3.5 that the output of the OSCILLATOR BANK as a whole has its region of maximum amplitude around a period of 600 ms. This centring on 600 ms arises from (1) the way that oscillators nearer to OSCILLATOR 6 are assumed to have higher amplitudes

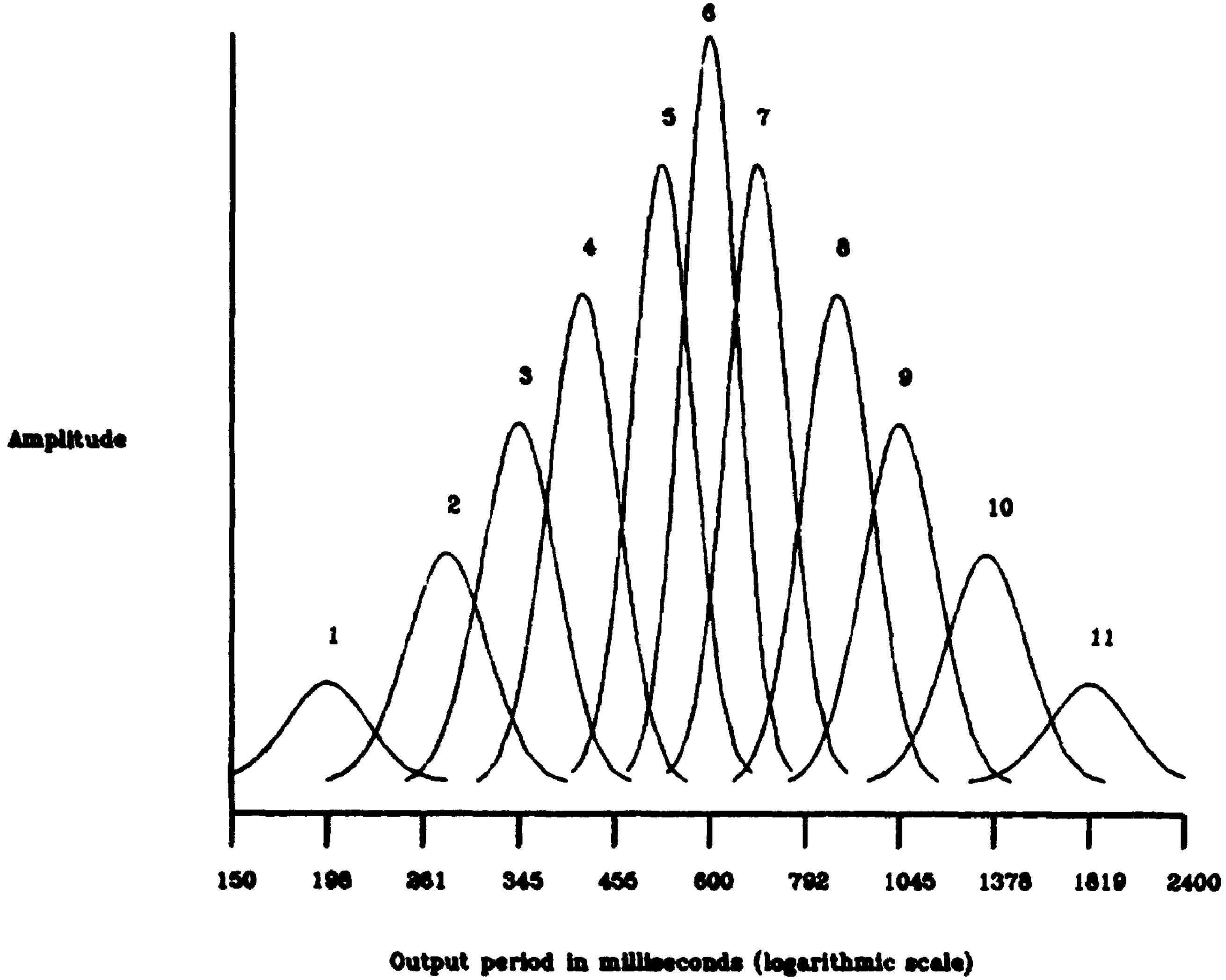


Figure 3.5. Output period ranges of oscillators shown as bell curves. Numerals next to curves refer to the oscillator number in the OSCILLATOR BANK.

than those further from OSCILLATOR 6; (2) the way that oscillators are more densely packed in the middle of the overall period range than they are at the extremes. For example, from Table I, it is evident that a period of, say, 650 ms can be produced by OSCILLATORS 5, 6, 7 and 8. At this period OSCILLATORS 5 and 8 can only produce outputs of a fairly meagre amplitude, nonetheless the combination of the four oscillators results in a high aggregate amplitude. Conversely, a period of, say, 190 ms can only be produced by OSCILLATOR 1. Although OSCILLATOR 1 is near the middle of its range with a period of 190 ms, even at its maximum, this oscillator can only output a fraction of the amplitude produced by OSCILLATOR 6. Therefore the overall output at a period of 190 ms will have a low amplitude.

As mentioned above, the oscillators in the OSCILLATOR BANK can be driven by pulses from the IEIT. In other words, the oscillators can adapt their output periods to the time intervals between IEIT pulses, and thus to IOIs presented in a stimulus. In this way the auditory system tunes neural timing mechanisms to events in the environment. How the IEIT pulses reach the OSCILLATOR BANK depends on the state of the oscillators. An oscillator is presumed to occupy one of three states: *pre-activation*, *activation* and *entrainment*.

Pre-activation. An oscillator is in the pre-activation state before it has received any input pulses. In this state, oscillators oscillate at random periods and phase relative to one another. The pre-activation period is likely to be near the oscillator's mean period. For the purposes of illustration in this model, the assumed pre-activation period for each oscillator will be the corresponding mean period shown in Table I. The amplitude of these random oscillations is very low (hereafter called the "hum level"), and the human "owner" of the oscillators is not consciously aware of them. Oscillation in the pre-activation stage is similar to the "resting period" of an adaptive oscillator used in a recently developed computer model of entrainment (Large, McAuley and Kolen 1993). The random, imperceptible activity of the OSCILLATOR BANK in the pre-activation state is exemplified in Figure 3.6.

In the pre-activation state, the input to each oscillator is controlled by an input filter (refer again to Figure 3.3). Each input filter is a band-pass type whose response range matches its oscillator's output-period range, that is, the ranges given in Table I. The response ranges of the input filters for the pre-activation state are rendered graphically in Figure 3.7. IEIT pulses are passed through these filters to the input of the oscillators only if the inter-pulse time interval (the stimulus IOI) lies within the filter's range. Since it takes two pulses to make an IOI, it is always the *second*, concluding pulse that is passed through a filter to the corresponding oscillator.

Activation. If an IOI matches the input filter's response range, then the IEIT pulse that concludes the IOI causes activation of the oscillator. Thus if two IEIT pulses a and b are separated by, say, 500 ms, pulse b will reach OSCILLATORS 4, 5 and 6 and activate those oscillators. Activation of any oscillator (OSCILLATOR X) entails the following (illustrated in Figure 3.8):

1. OSCILLATOR X is pulled up from hum level to a markedly higher (maximum) amplitude

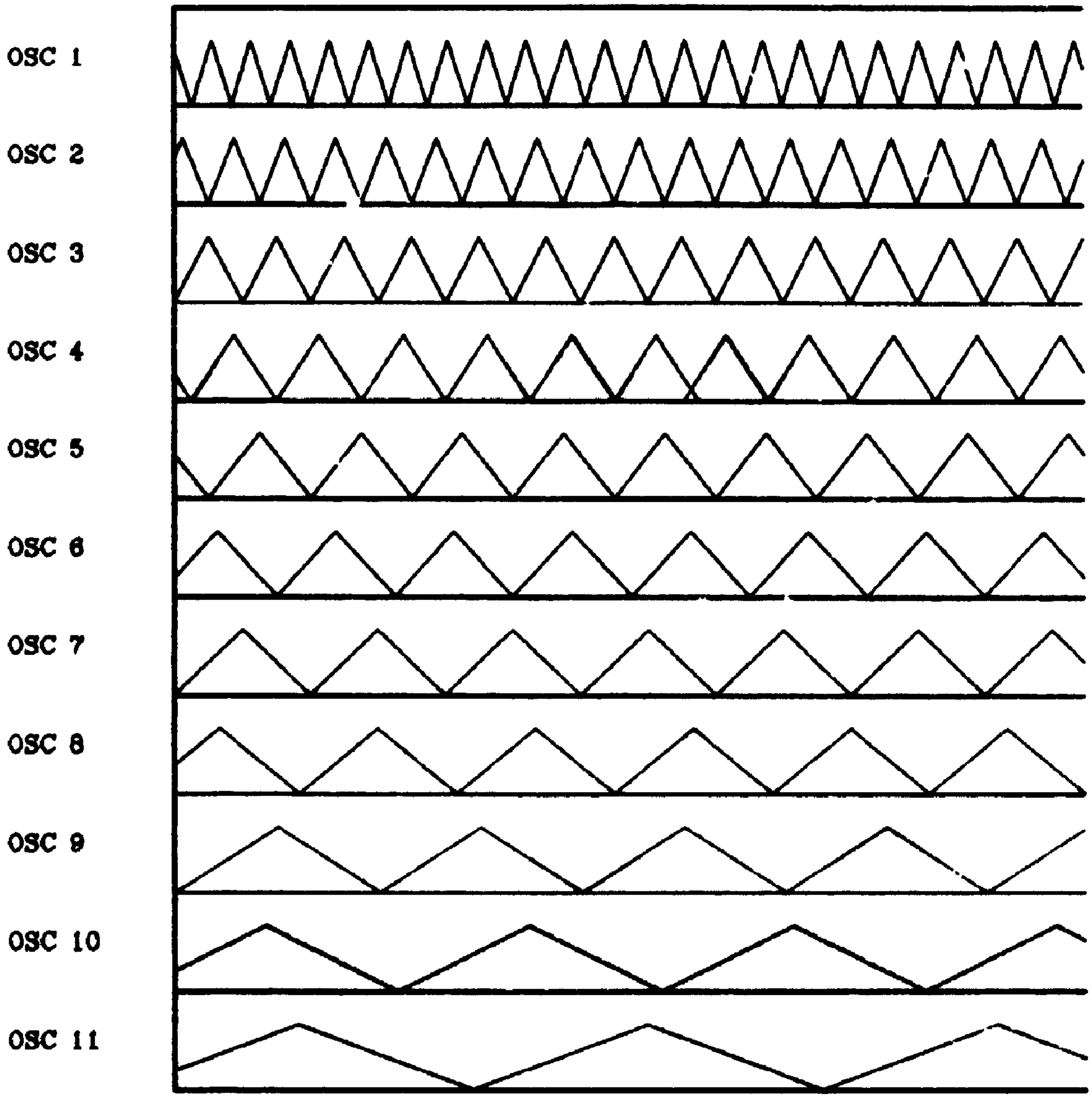


Figure 3.6. Output of OSCILLATOR BANK in the pre-activation state. Oscillators produce waveforms of random period and phase, at a low amplitude

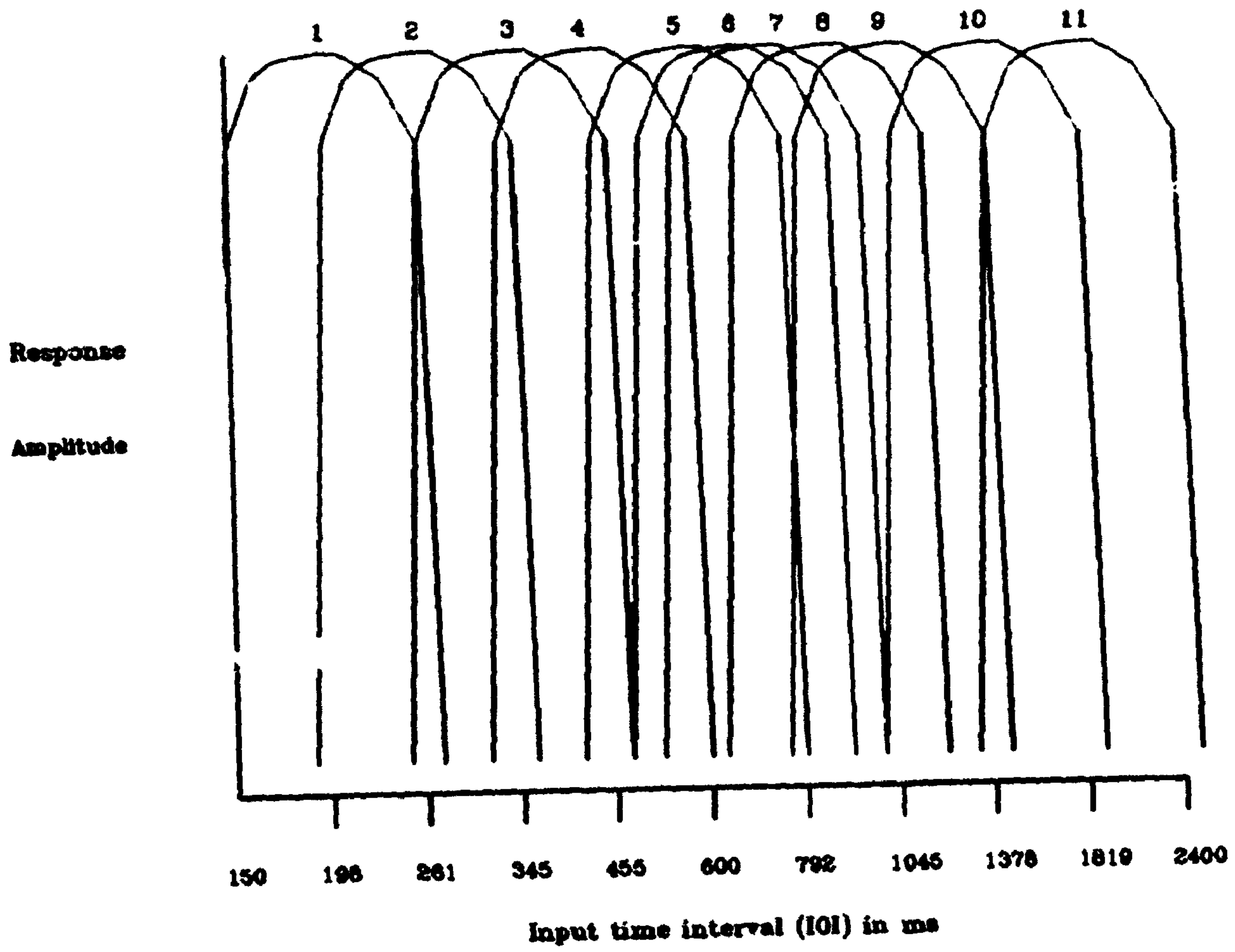


Figure 3.7. Response ranges of input filters inthe OSCILLATOR BANK. Numerals next to curves indicate the number of the oscilator whose input filter has the period response represented by the curve.

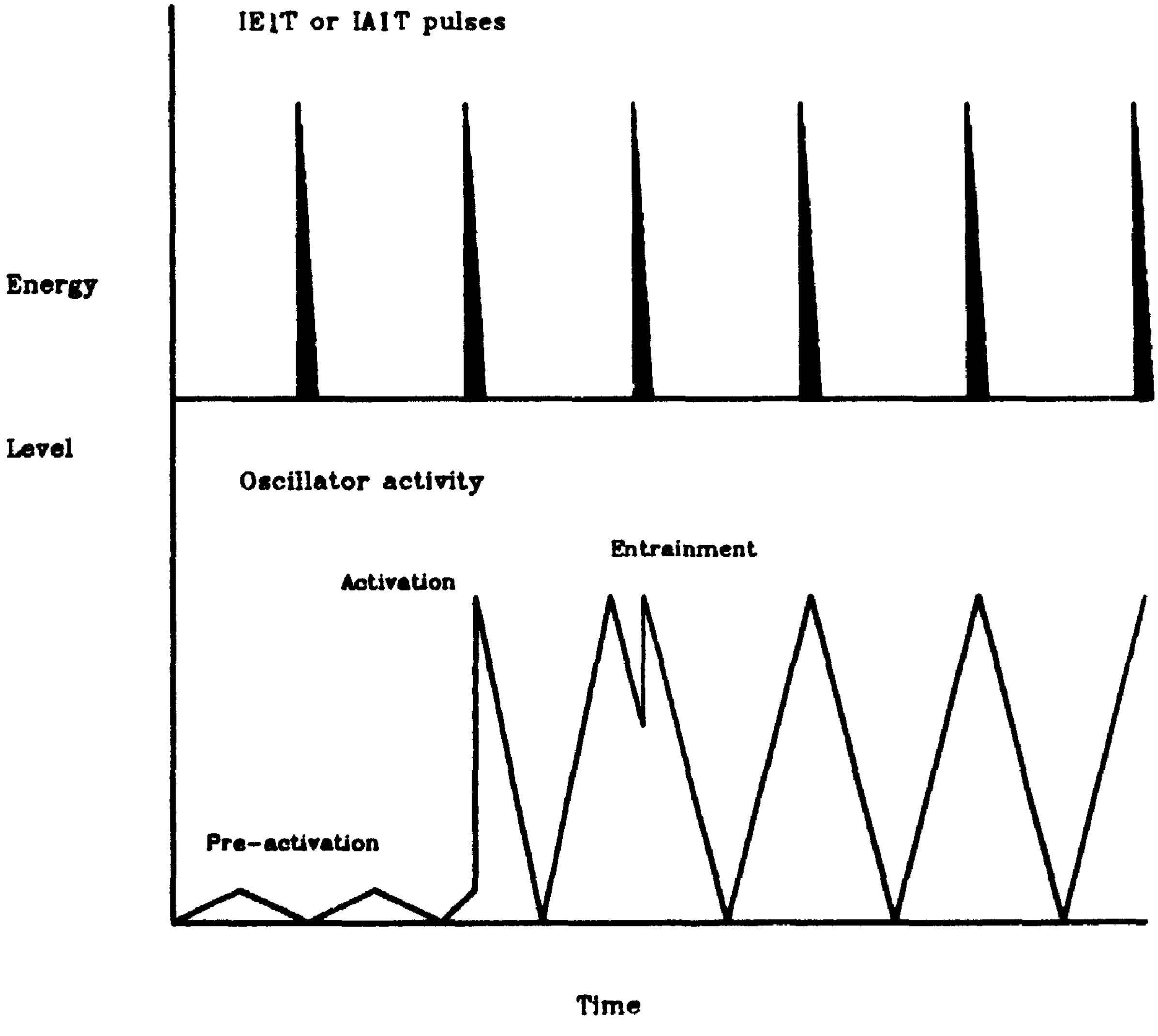


Figure 3.8. Activation and entrainment of a typical oscillator (OSCILLATOR X).

2. consequently, for the first time, the owner could now become consciously aware of the activity of OSCILLATOR X
3. owing to the amplitude change, the phase of OSCILLATOR X is reset, so that the start of its cycle (maximum energy level) coincides with the onset of the IEIT pulse
4. despite phase resetting, for the time being, OSCILLATOR X still oscillates at the random period at which it had oscillated in the pre-activation state
5. OSCILLATOR X's input filter is now by-passed, so that the oscillator is henceforth connected directly to the output of the IEIT.

Once activated, OSCILLATOR X continues oscillating at some randomly chosen period (its mean period). If there are no subsequent stimulus IOIs to which the oscillator can entrain, OSCILLATOR X will decay back to hum level after an arbitrary time span (for the sake of argument, 2 seconds), and at that point the input filter is reconnected. If there are subsequent stimulus IOIs, entrainment may take place.

Entrainment. If OSCILLATOR X has been activated by the first IOI (second onset), it may then entrain to the IOI between the second and third onsets. When entrainment takes place two changes occur in OSCILLATOR X: (1) the oscillator is again phase reset by the IEIT pulse that concludes the second time interval; (2) the oscillator's period is pulled towards the value of the second time interval. Again, Figure 3.8 shows the general process of entrainment following activation.

However, once OSCILLATOR X has been activated, for phase resetting (and therefore entrainment) to take place, the oscillator's instantaneous energy level at the point in time of resetting must be above a particular threshold level. The threshold energy level may be standardized as a fraction:

$$I_{tr} = EL_{tr}/EL_{max} \tag{3.1}$$

where I_{tr} is the "threshold index", EL_{tr} is the threshold energy level for phase resetting, and EL_{max} is the maximum energy level (the apex of the ramp waveform) produced by OSCILLATOR X at the period that was established just prior to phase resetting. Referring back to Figure 3.5, it can be seen that the absolute value of EL_{max} will differ among oscillators, and for one oscillator across its period range. For the purposes of this model, I_{tr} is held constant at .6, that is, EL_{tr} must be equal to or greater than 60% of EL_{max} for phase resetting to take place. Examples of the different ways in which an IEIT pulse might coincide with the ramp waveform of an oscillator are given in Figure 3.9.

It should be evident from Figure 3.9 that the lower the instantaneous energy level of the oscillator waveform at the point in time when the IEIT pulse occurs (EL_i), the greater the asynchrony between the oscillator's period and the stimulus IOI. In fact, any instantaneous energy level, EL_i, expressed as a fraction of EL_{max} may be translated into a corresponding "asynchrony value." An asynchrony value is defined as the degree to which the IEIT pulse is early or late in relation to the currently established oscillator period. With a ramp waveform, the formula that converts relative energy level to relative asynchrony is as follows:

$$AS_{rel} = .5(1 - EL_i/EL_{max}) \tag{3.2}$$

where AS_{rel} is the relative asynchrony value expressed as a fraction of the oscillator's period, and EL_i and EL_{max} are as previously defined (a

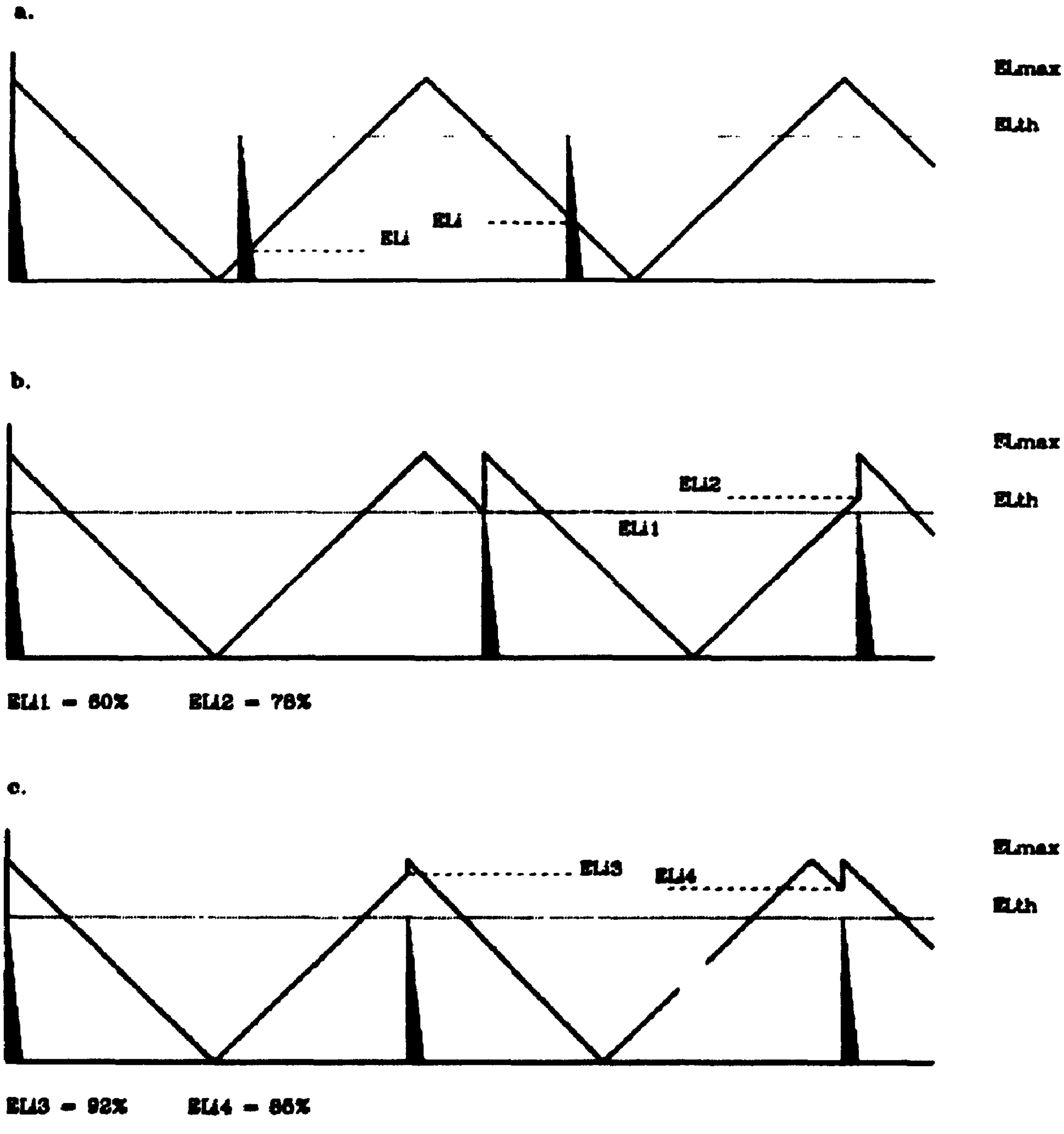


Figure 3.9. Three different types of coincidence of IEIT pulses and oscillator activity. Figure 3.9a shows failure to reset. Figure 3.9b shows phase resetting with with marked asynchrony. Figure 3.9c shows phase resetting with mild asynchrony.

geometrical derivation of this formula may be found in Appendix B). Since the threshold index, $I_{..}$, as previously defined, is itself a fraction of EL_{-3x}, the relative asynchrony allowed by this index may be calculated:

$$AS_{re} = .5(1 - .6) = .2$$

Therefore with a threshold index of .6, the IEIT pulse may be asynchronous with the beginning of the oscillator's cycle by up to +/- 20% of the oscillators established period for phase resetting still to take place. This range of asynchronies within which phase resetting may take place is termed the "acceptable asynchrony range."

If phase resetting can occur, the number of oscillator cycles needed to entrain to an approximately isochronous stimulus depends on the degree of difference between the oscillator's currently established period and the to-be-entrained period. For the purposes of the present model, it is assumed that, with one phase resetting, an oscillator period can be pulled towards the to-be-entrained period by up to +/- 20% of its currently established period. For example, if an oscillator with a currently established period of 650 milliseconds is phase reset by an an isochronous set of IOIs of 505 ms, the oscillator will need two phase resettings to entrain to 505 ms since the most it can be pulled by one phase resetting will be 20% x 650 or 130 ms (i.e. down to 650 - 130 or 520 ms).

For a further example, take OSCILLATOR 5. Suppose it is exposed to two IOIs of 500 ms each, bounded by three event onsets, *a*, *b* and *c*. Suppose also that this oscillator has randomly chosen a period of 520 ms in the pre-activation state (OSCILLATOR 5's mean period). The first IOI of 500 ms (*a*-*b*) is passed through OSCILLATOR 5's input filter causing activation. Thus at event onset *b*, OSCILLATOR 5 is pulled up to an activated amplitude level (turned on), its phase is reset with *b*, its filter is bypassed and it begins a new cycle, but still at its initial period of 520 ms. Then, 500 ms after *b* at event onset *c*, a new IEIT pulse occurs, but this is 20 ms earlier than the time-point when OSCILLATOR 5 is next due to reach EL_{max} again. Whether the oscillator can phase reset with this asynchronous IEIT pulse, depends on the oscillator's instantaneous energy level at that point in time. Solving equation 3.2 for EL_i/EL_{max} one gets:

$$EL_i/EL_{max} = 1 - 2(AS_{rel}) \tag{3.3}$$

In this example $AS_{rel} = 20/520 = .0385$, therefore

$$El_i/EL_{max} = 1 - 2(.0385) = .923.$$

In other words, when the IEIT pulse occurs, OSCILLATOR 5 has reached roughly 92% of its maximum energy level. This value is easily above the threshold index of .6 (60%), so phase resetting can readily take place. With the phase resetting at event *c*, OSCILLATOR 5 is now entrained to 500 ms, a period that it will produce in the immediate future should no other phase resetting take place in that oscillator. Should subsequent events present a novel time interval that is nonetheless within the acceptable asynchrony range of the entrained period, and within the oscillator's output-period range, this oscillator will phase reset and be pulled toward the new time interval.

Generally then, the sequence of events for the activation and entrainment of OSCILLATOR X may be summarized as in Table II. Note that the left-hand column represents stimulus events *a*, *b*, *c*, etc. in sequential order. From Table II it can be seen that, generally, OSCILLATOR X can entrain to a stimulus period after only three or four stimulus onsets have occurred, and re-entrain should mildly different IOIs occur subsequently. Thus, inaccuracies or expressive deviations in performance timing do not prevent this system from tuning in to time intervals inherent in a musical stimulus.

TABLE II

CORRESPONDENCE BETWEEN STIMULUS AND OSCILLATOR X

Stimulus Event	Activity of OSCILLATOR X
No stimulus	Oscillator in *pre-activation* state: randomly chosen period and phase, hum-level amplitude.
a	Ditto
b	If IOI between *a* and *b* is within OSCILLATOR X's input filter range, oscillator *activated*: phase reset, amplitude high, period same as before. Oscillator's input filter is by-passed.
c	If IOI between *b* and *c* is within the acceptable asynchrony range, oscillator *entrained*: phase reset, period pulled towards value of IOI between *b* and *c*. Amplitude depends on where entrained period lies in oscillator's output range.
. . . .	
m	
n	If IOI between *m* and *n* is novel, yet within the acceptable asynchrony range, oscillator *re-entrained*: phase reset, period pulled towards value of IOI between *m* and *n*. Amplitude depends on where re-entrained period lies in oscillator's output range.

All eleven oscillators follow the same basic operating principles. Since the oscillators' input filters respond to different yet overlapping time-interval ranges, some oscillators are activated by the same stimulus event, and others must wait for another more suitable IOI for activation. Nevertheless, it is theoretically possible that, after only several onsets in a durationally rich stimulus rhythm, all eleven oscillators could be oscillating simultaneously.

Given the right conditions, all oscillators, once activated, can re-entrain to new IOIs in this mechanistic fashion *ad infinitum*. The human "owner" of the OSCILLATOR BANK will only become consciously aware of phase resetting if the discrepancy between the already-entrained period and the novel period is sufficiently large. For the purposes of this model it is proposed that, provided that AS_{rel} is small enough for phase resetting to occur in the first place (within the acceptable asynchrony range), the unweighted probability of a listener becoming aware of phase resetting in one oscillator is as follows: the currently-established maximum energy level (EL_{max}) minus the instantaneous energy level (EL_i) of the oscillator's waveform at the instant of the resetting pulse, expressed as a fraction. In other words,

(3.4) $P_R = 1 - (EL_i/EL_{max})$, provided EL_i/EL_{max} >/= I_{thr},

where P_R is the unweighted individual probability of a phase reset being noticed, and EL_i, EL_{max} and I_{thr} are as previously defined.

The probability value P_R is said to be "unweighted" and "individual" because this model proposes that the overall probability of a particular anomalous onset being noticed is the summation of all the weighted probabilities of the individual oscillators that must phase reset as a result of the anomalous onset. This summation may be expressed as follows:

(3.5) $$OP_R = \Sigma wP_R$$

where OP_R is the overall probability of an anomalous onset being detected, P_R is the unweighted individual probability of detecting a resetting in one oscillator as defined above. The variable w is the weight ascribed to each oscillator according to the following equation:

(3.6) $$w = 1 - (|x - 6| + 1)/10$$

where x is the number of the oscillator. From equation 3.6, the weights (values of w) given to oscillators in the OSCILLATOR BANK are: OSCILLATOR 6 = .9, OSCILLATORS 5 and 7 = .8, OSCILLATORS 4 and 8 = .7, OSCILLATORS 3 and 9 = .6, OSCILLATORS 2 and 10 = .5, OSCILLATORS 1 and 11 = .4. These weights are somewhat arbitrary, however, they reflect the preferred tempo range identified in the literature in Chapter 2. Two corollaries of equations 3.5 and 3.6, then, are as follows: (1) the more oscillators that are reset by an anomalous onset the greater the value of OP_R; (2) the closer those oscillators are to OSCILLATOR 6 the greater the value of OP_R.

To exemplify the calculation of the overall probability of a listener detecting an anomolous onset, consider the situation where an IOI of 630 ms is inserted in an otherwise isochronous pattern with a repeated IOI of 550 ms. In such a case, OSCILLATORS 4, 5, 6 and 7 would be able to entrain to the initial IOI of 550 ms. Of these, OSCILLATORS 5, 6 and 7 would be able to re-entrain to the IOI of 630 ms. In addition, OSCILLATOR 8 would be activated by the IOI of 630 ms, but would not yet be entrained so it may be discounted from the ensuing discussion.

Firstly, the relative asynchrony of the 603-ms IOI can be computed:

$$AS_{re} = (630 - 550)/550 = .14545$$

Secondly, substituting this value in equation 3.3 one gets:

$$EL/EL_{max} = 1 - 2(.14545) = .70909,$$

that is, the synchronized waveforms of OSCILLATORS 5, 6 and 7 are all at 71% of their maximum energy level at the instant of phase resetting at the completion of the 603-ms IOI. Thirdly, from equation 3.4, the unweighted probability of such an asynchrony being noticed is

$$P_R = 1 - (.70909) = .29091$$

Fourthly, substituting this P_R value into equation 3.5, with the appropriate weights for OSCILLATORS 5, 6 and 7, one gets:

$$OP_R = .8(.29091) + .9(.29091) + .8(.29091) = .72727.$$

In other words, according to this model, there is a roughly 73% overall probability of this anomalous onset being detected in the OSCILLATOR BANK.

The foregoing examples of phase resetting, and the corresponding probabilities of the phase resetting being noticed, have so far dealt only with the processing of time intervals on one metric level. In other words, although anomalous IOIs have been introduced, they have not been so different from the entrained period that they establish a second integrally related set of oscillations. If, however, subdivisions or groupings on the beat time interval were introduced in the stimulus, then several levels of a time-interval hierarchy would be entrained.

For instance, if IOIs of 250 or 1000 ms were introduced into the 500-ms example given above, then OSCILLATORS 1 and 2 would entrain to 250 ms, OSCILLATORS 4, 5 and 6 would entrain to 500 ms and OSCILLATOR 8 would entrain to 1000 ms. In terms of Jones and Boltz's (1989) model, the result would be three levels of a binary hierarchy entrained in the OSCILLATOR BANK. The current model proposes that, in such a case, the owner of the OSCILLATOR BANK could attend to any one of the three levels, but not more than one concurrently. The overall probability of a phase resetting being noticed with a metrically multi-leveled stimulus, then, would depend on which level of activity of the OSCILLATOR BANK the owner attends to at the point of phase resetting. This situation will be discussed further in formulating the experimental hypotheses in Chapter 4.

As a final, more general example of the workings of the OSCILLATOR BANK, Figure 3.10 illustrates the activity of OSCILLATORS in response to the two rhythms (the 6/8 and 2/4 contexts) used by Clarke in his (1987a) perceptual experiment (see Figure 2.3 in the previous chapter).

The Short Auditory Store

Having outlined the hypothetical workings of the IEIT and its interaction with the OSCILLATOR BANK, one can now turn to the SHORT AUDITORY STORE, or SAS, and the INTER-ACCENT INTERVAL TIMER, or IAIT.

In this model, the SAS is responsible for the concatenation of stimulus events that are physically separated in time. For this reason, an epiphenomenon of its operation is the psychological present in the

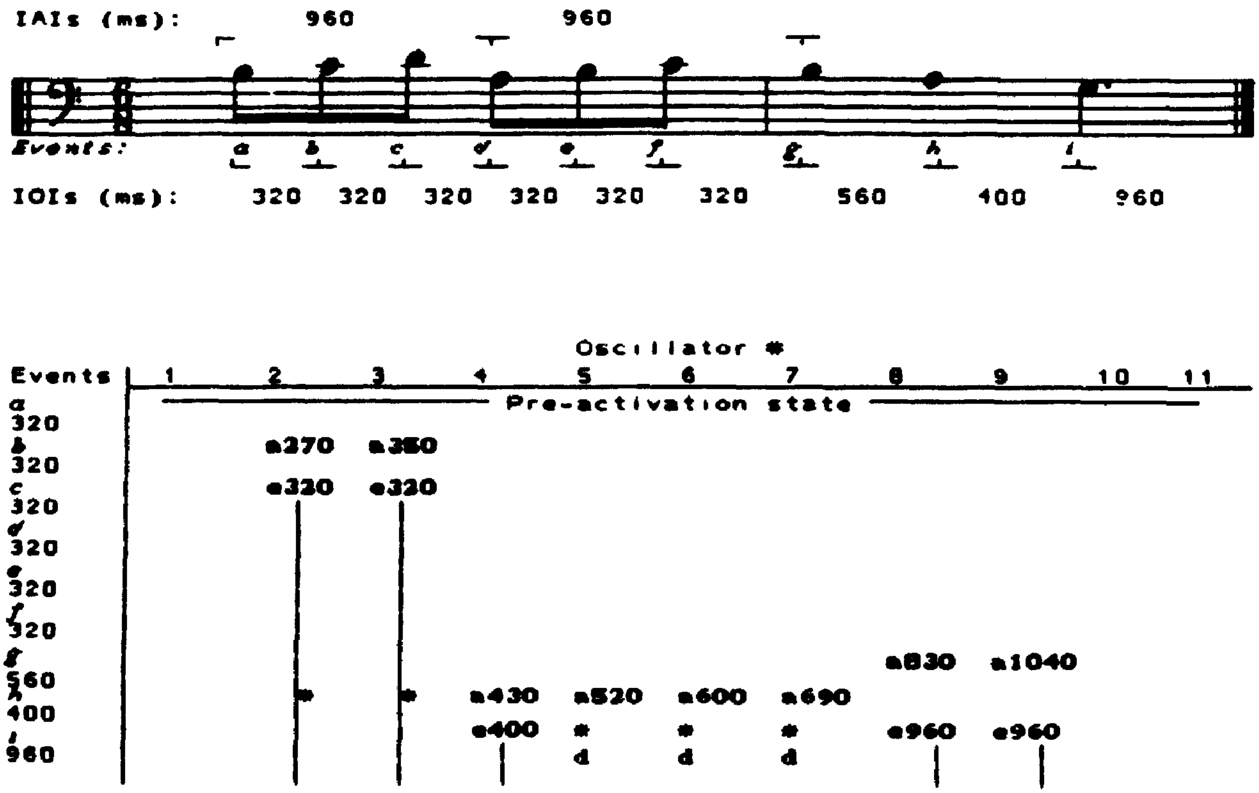


Figure 3.10a. OSCILLATOR BANK response to Clarke's (1987a) rhythm number 5 in the 6/8 context. The head of the figure shows the rhythm in terms of IOIs and IAIs. The main body of the figure shows the activity of oscillators over time, with time going vertically down the page. In both this figure and Figure 3.10b the following abbreviations are used: ax = activation at a period of x ms (oscillator's mean period); ey = entrainment to a period of y ms; rz = re-entrainment to a period of z ms; * = failure to phase reset at a particular event; d = decay owing to failure to entrain.

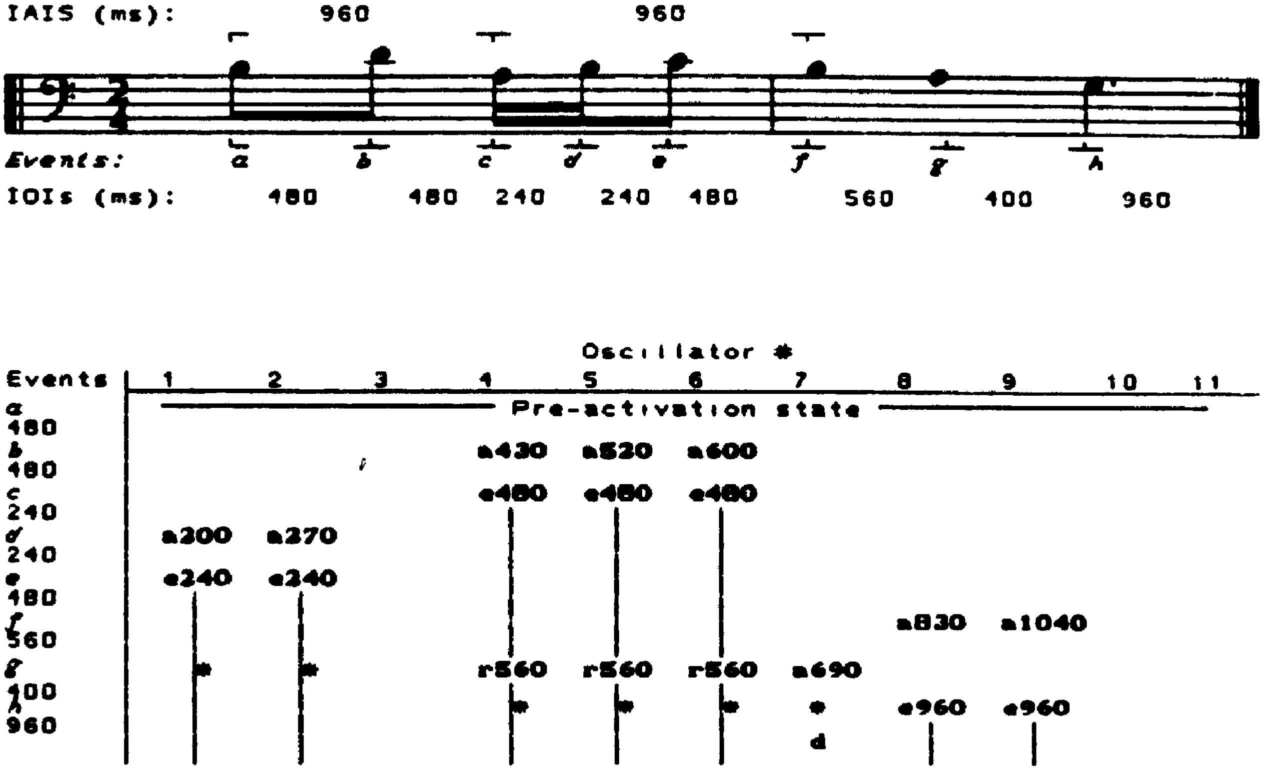


Figure 3.10b. OSCILLATOR BANK response to Clarke's (1987a) rhythm number 5 in the 2/4 context. The head of the figure shows the rhythm in terms of IOIs and IAIs. The main body of the figure shows the activity of oscillators over time, with time going vertically down the page. The abbreviations are the same as in Figure 3.10a.

auditory modality. In the system proposed here, the SAS is intended to model the observed behaviours of both echoic memory and STM in the auditory modality. The particular model of STM that is taken as a starting point is Frick's (1988) model of short-term auditory memory.

The hypothetical workings of the SAS are depicted by analogy in Figure 3.11. Instead of Frick's single variable-speed tape recorder, one might imagine a variable-speed sampler. This variable-speed sampler model differs from Frick's tape-recorder analogy, not only in the recording medium used as an analogy, but also in the number of stages of recording envisaged. In fact, what is here proposed is a multi-stage, variable-speed sampler consisting of several SAMPLING BUFFERS. In Figure 3.11, SAMPLING BUFFER A is the input to the SAS, and, as such, models the workings of echoic memory or precategorical acoustic storage. A "chunk" of the input stimulus is first recorded in SAMPLING BUFFER A with most of its acoustic detail intact. SAMPLING BUFFER A has a maximum storage capacity of between 2 and 3 seconds, however, if the stimulus is complex and the listener wishes to retain acoustic detail, then SAMPLING BUFFER A will have to sample at a faster rate to enable a recording of the desired fidelity to be made. Evidently if SAMPLING BUFFER A samples at a fast rate, its storage capacity will be used up quickly. The result is a trade-off between recording length and recording fidelity.

Once SAMPLING BUFFER A is full, it begins to record a new chunk of the stimulus erasing the old recording. So that the old sample is not lost, however, as its space in SAMPLING BUFFER A is taken up by a new chunk of stimulus, the old sample is shunted to the next SAMPLING

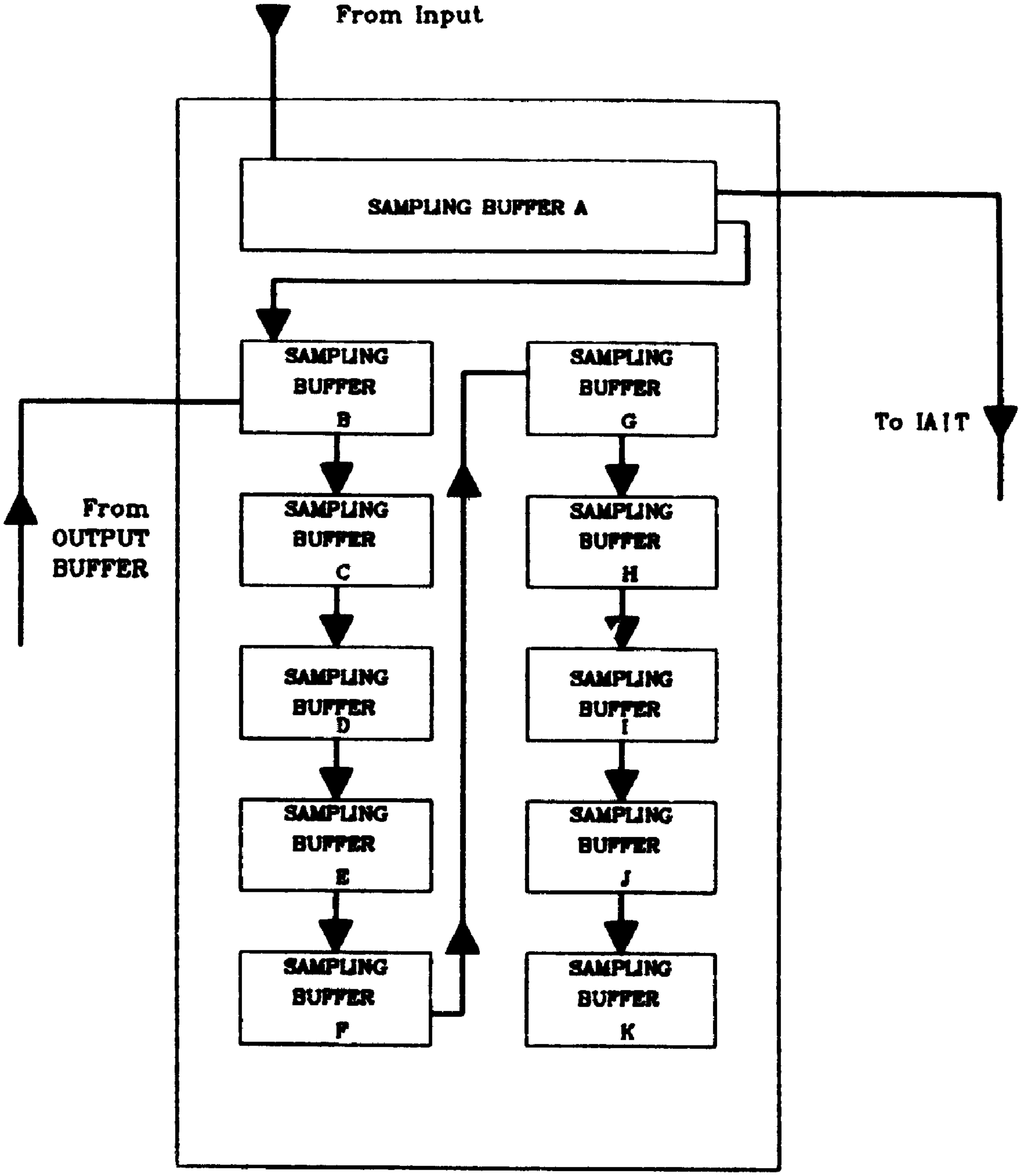


Figure 3.11. Detailed illustration of the SHORT AUDITORY STORE (SAS).

BUFFER "down the line," SAMPLING BUFFER B. Unfortunately, since this process entails rerecording, one of the results of shunting is a loss in fidelity. The contents of SAMPLING BUFFER B will therefore be a distorted, impoverished version of the signal originally encoded in SAMPLING BUFFER A. Once SAMPLING BUFFER A is full again, it begins a new chunk, its current contents are shunted to SAMPLING BUFFER B, and the contents of SAMPLING BUFFER B are in turn shunted to SAMPLING BUFFER C. The process continues: with each subsequent chunk of stimulus that is recorded in SAMPLING BUFFER A, older chunks are shunted further down the line from SAMPLING BUFFER to SAMPLING BUFFER. With each shunting (rerecording) more and more acoustic detail is lost until the recorded signal has deteriorated completely.

Given the process described above, the most vivid recording is found in SAMPLING BUFFER A, the contents of which comprise the present event and events that have just occurred. The contents of SAMPLING BUFFER A are instantly accessible, either as a complete grouping or, possibly, as individual items. SAMPLING BUFFER A is therefore posited as the mechanism responsible for the phenomenon of the psychological present. The contents of SAMPLING BUFFERS B, C, D, and so on down the line, are less and less vivid, and may be several seconds old. Access to these buffers may not be gained instantly, and the contents are less manipulable than those of SAMPLING BUFFER A. In addition, as will be described later, the contents of these later buffers are no longer pure signal, but are mixed with other categorical information about the signal. SAMPLING BUFFERS B, C, D, etc. are therefore included to model auditory STM. The time taken for signal deterioration due to shunting models decay time for STM. Since SAMPLING BUFFER A models echoic memory, and SAMPLING BUFFERS B, C, D etc. model auditory short-term memory, comparison between present perception and short-term memory would be carried out by comparing the contents of SAMPLING BUFFER A with the contents of the other SAMPLING BUFFERS.

A second consequence of shunting a signal trace down the line is a distortion of the absolute time intervals of the original stimulus. Since the most vivid and most recent recording is always held by SAMPLING BUFFER A, this particular buffer is accorded the most importance by the perceptual system. The further down the line a SAMPLING BUFFER is, the less recent and therefore the less sharply-etched are its contents. Therefore, in the interests of cognitive economy, buffers further down the line are apportioned less and less storage medium. Since each subsequent buffer has less and less space to record the time-span of the signal passed down to it, the further down the line a buffer is, the more it must compress a signal trace in the temporal domain.

Another way of viewing this process of temporal compression is to recognize that, owing to decay of the pure signal trace, the contents of buffers further down the line will have a greater proportion of abstract information compared to raw signal data. This aspect of the model is similar to certain models of speech processing in psycholinguistics (e.g., Beaugrande 1985). In such models, "shallower" levels of processing (e.g., phoneme processing) are assumed to take place primarily in the first couple of seconds of a sound's presentation, although, for example, "long-term memory can retain some traces of the surface, e.g., the recollection of whether an item was presented visually or acoustically." (Beaugrande 1985, 166). Thus, if the contents of later buffers are increasingly more abstract representations of the signal trace, then temporal information must be collapsed or compressed in some fashion in these buffers. If memories of, say, 15 seconds old were still intact with all their acoustical (and therefore temporal) detail, they would take as long to "play back" in memory as they took to be presented in the original stimulus. However, according to the short-term memory literature cited in Chapter 2, subjects do not require the full time-span of the original stimulus to perform, say, a pattern comparison task between memory and perception. Therefore temporal compression seems to be a normal (and desirable) part of encoding in the SAS.

For the purposes of the types of stimuli considered in this dissertation, temporal compression in the SAS has one important ramification: Events stored in sampling buffers far down the line, SAMPLING BUFFERS G or H, for example, will seem faster than they originally sounded in perception. Put another way, if the system is presented with a perfectly isochronous rhythm, the present perception of that rhythm will seem slower than a several-second-old memory trace of the same rhythm. Therefore, if the listener compares the contents of SAMPLING BUFFERS G or H (his/her STM) with those of SAMPLING BUFFER A (his/her current perception), the rhythm will seem to have slowed down even though it has remained perfectly isochronous. In a second situation where the rhythm has, in fact, decelerated, any comparison between present perception and memory will overemphasize that deceleration. Conversely, in a third situation where the rhythm has accelerated, comparison between memory and perception will annul some of the difference between the presently perceived tempo and the tempo as stored in STM.

More generally, this analogical model parsimoniously models a great deal of the literature on precategorical storage and auditory memory. Most importantly perhaps, a distinction is made between the concepts of storage capacity and decay rate. The volume of echoic memory is equivalent to the amount of stimulus that can be recorded on SAMPLING BUFFER A. The decay rate depends on the amount of acoustical detail that is lost over time with each re-recording. In addition, the wide range of durations in auditory memory research noted by Cowan and others can be accounted for: both the storage capacity and the decay rate of auditory memory will vary depending on what a listener must attempt to hold in memory, i.e., what sort of recording fidelity is required. For this reason, then, the nature and difficulty of a subject's task in an experiment will determine the memory trace's volume and decay rate.

Measuring Time Intervals between Accents

The IAIT measures the time intervals that elapse between successive accents, a task which is considerably more complex than measuring simple time intervals between events. In measuring inter-event time intervals the IEIT simply waits for the next onset, and at that moment is able to measure the inter-event time interval. In contrast, in many examples of the accent types listed in the definition of accent in Chapter 1, the listener logically cannot know that an event is accented with respect to its neighbours until some time after the onset of the event has elapsed. Examples of what may be called *retrospectively assigned accents* are given in Figure 3.12.

In Figure 3.12a the fact that the first note is dynamically stressed with respect to the notes that follow cannot be known until at least the second tone has sounded. In Figure 3.12b the fact that the note *c* is longer than its two predecessors, and therefore enjoys durational accent, cannot be known by the listener until at least one eighth-note's duration after the onset of *c*. In Figure 3.12c the fact that the C is melodically prominent by virtue of its position at a point of contour change (Thomassen 1982) cannot be known until the following, lower note has sounded. In all three cases the information necessary to ascribe accent to particular events is not available at the time the onsets of those events occur. It is intriguing that phenomenologically all three examples sound as if the accented event receives its accent at the point in time of its onset.

The observation that listeners seem to hear events as accented at the very point in time when they occur, although they cannot at that point know if they are accented or not, is compelling evidence for some type of temporally integrative process in auditory perception, for example SAMPLING BUFFER A. In the current model, then, the IAIT measures the time intervals between accented events, not as they occur in time, but in a stored format in the SAS, specifically in SAMPLING BUFFER A.

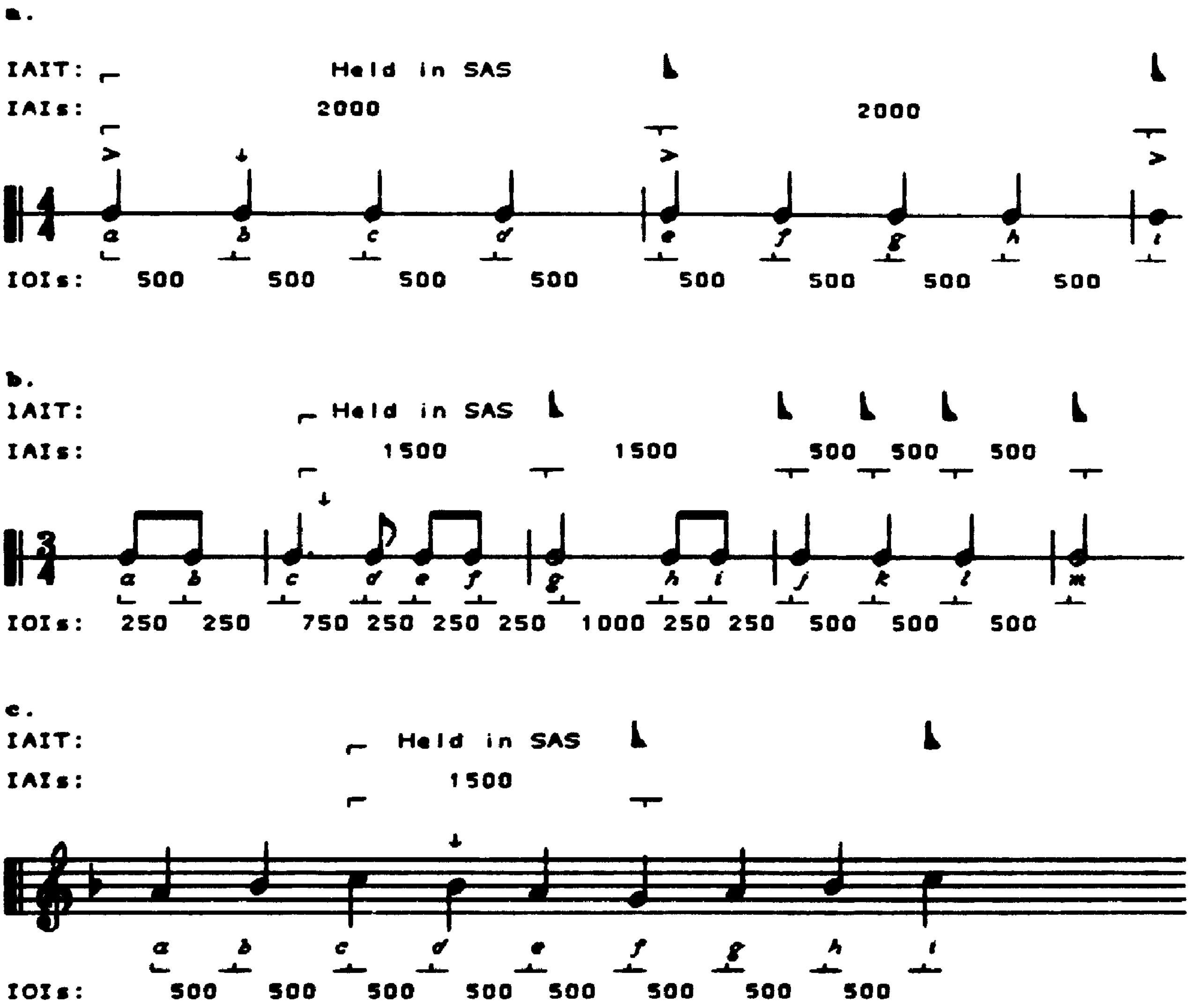


Figure 3.12. Retrospectively assigned accents. In each case the point in time where the accent is retrospectively assigned is marked with an arrow. The pulse activity of the IAIT in response to each stimulus is shown above each stave. Figure 3.12a exemplifies dynamic stress. Figure 3.12b exemplifies durational accent. Figure 3.12c exemplifies melodic or pitch-related accent.

One might illustrate the activity of the IAIT by returning to the examples in Figure 3.12. In the case of the dynamic accent shown in Figure 3.12a, at the onset of event *b* the IAIT realizes that *a* was louder. The IAIT therefore places a marker on event *a* in the SAS, and then waits for a subsequent event of a similar dynamic level. When the next accented event occurs (*e*), the IAIT generates a pulse. Next, the IAIT retrieves the time interval *e-a* from the SAS and generates a second pulse at *e-a* (2000 ms) after *e*.

The output pulses from the IAIT are passed to the OSCILLATOR BANK. The OSCILLATOR BANK handles pulses from the IAIT exactly as it does pulses from the IEIT. In other words, the first inter-pulse time interval received by a given oscillator causes activation, and the second inter-pulse time interval received by the same oscillator brings about entrainment. Generally, however, since IAIs are longer than (multiples of) IOIs, it will normally be the slower oscillators (OSCILLATORS 9, 10 AND 11) that entrain to the time intervals between IAIT pulses.

Returning to the examples of retrospectively assigned accent shown in Figure 3.12, in the case of the durational accent shown in Figure 3.12b., at 250 ms after the onset of *event c* the IAIT realizes that the IOI following event *c* is longer than the two previous events (*a* and *b*). The IAIT therefore places a marker on event *c* in the SAS and waits for a second event that is longer than the duration of *a* or *b*. When the next durationally accented event occurs at *g*, a pulse is generated by the IAIT. The IAIT retrieves the time interval *g-c* from the SAS and generates a second at *g-c* (1500 ms) after *g*. Again, these pulses are passed on directly to the OSCILLATOR BANK.

In the case of the melodic accent shown in Figure 3.12c., at the onset of event *d* the IAIT realizes that event *c* lies at a point of contour change relative to the surrounding notes (*a b* and *d*). The IAIT therefore places a marker on event *c* in the SAS and waits for a second event that also lies at a contour change. When the next melodically accented event occurs at *f*, the IAIT generates a pulse. The IAIT retrieves the time interval *f-c* from the SAS and generates a second pulse at *f-c* (1500 ms) after *f*. Again, these pulses are passed on directly to the OSCILLATOR BANK.

The stimuli exemplified in Figure 3.12 only present one type of accent at a time, a state of affairs that is simpler than many situations in real music. In this model it is hypothesized that the IAIT can attend to several dimensions of a musical stimulus simultaneously. Although the exact number of these dimensions may be determined by musical ability/training, the IAIT is able to generate pulses in response to dynamic, melodic and timbral accents sounding in synchrony or out-of-step with one another.

It should be stressed that the IAIT is not responsible for the listener's phenomenal experience of accented events. Its purpose is more modest: the IAIT simply extrapolates the time interval between accents and generates corresponding pulses that can be used as cues for establishing a metrical framework. The construction of a representation of meter is carried out by the next stage to be discussed, the METER GATING.

Establishing Metrical Structure

The output of the oscillators in the OSCILLATOR BANK, whether driven by IEIT pulses or IAIT pulses, together form the input to the METER GATING. The METER GATING may be thought of as a mechanism that sorts and sifts the various oscillator periods from the OSCILLATOR BANK and the inter-pulse time intervals from the IAIT. Two sub-stages of energy-level-sensitive gating are hypothesized, as illustrated in detail in Figure 3.13.

The first sub-stage of gating, what will be called "SQUARING," only lets an oscillator output waveform pass at the points in its cycle where its instantaneous energy level exceeds a certain threshold value. This threshold, similar in concept to the I_{th} described earlier, is here called the "squaring index," or I_{sq}, and is set for the purposes of this model at .6 of the oscillator's established maximum energy level (see Figure 3.14a). When the ramp waveforms produced by the oscillators in the OSCILLATOR BANK are subjected to SQUARING, the portions of the oscillator's cycle below the threshold energy level are, so to speak, "switched off." The resulting output waveform from the SQUARING sub-stage is a squared ramp waveform, as shown in Figure 3.14b. As can be seen by returning to Figure 3.13, each oscillator output passes through its own dedicated sub-stage of SQUARING.

The second sub-stage of METER GATING, what will be dubbed "NEATENING," takes the output of SQUARING as its input. NEATENING identifies the shortest (fastest) period to pass through SQUARING and takes this period as a trigger pulse. In other words, the (squared) output of the fastest oscillator to have been activated at any given

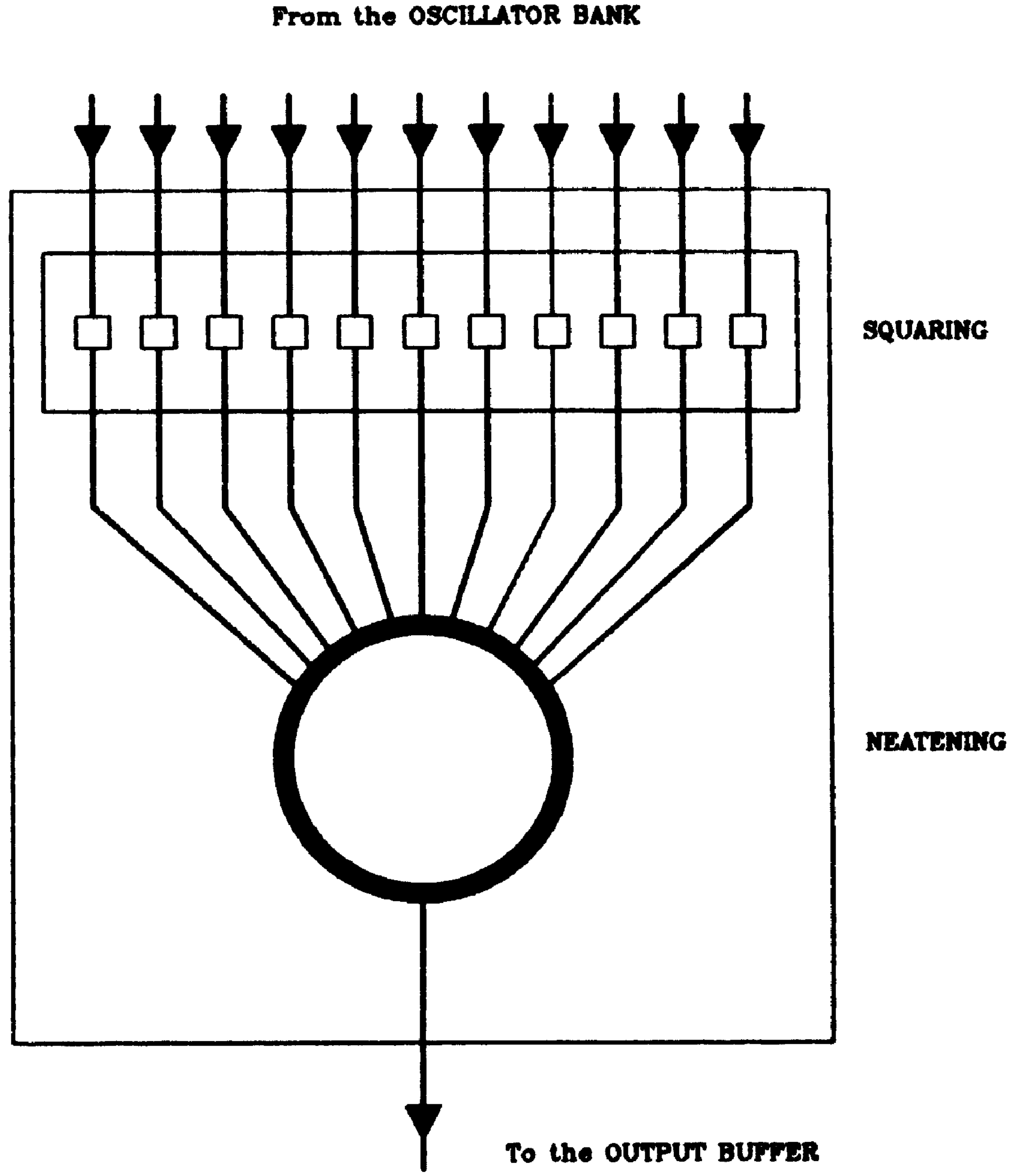


Figure 3.13. Detailed illustration of METER GATING.

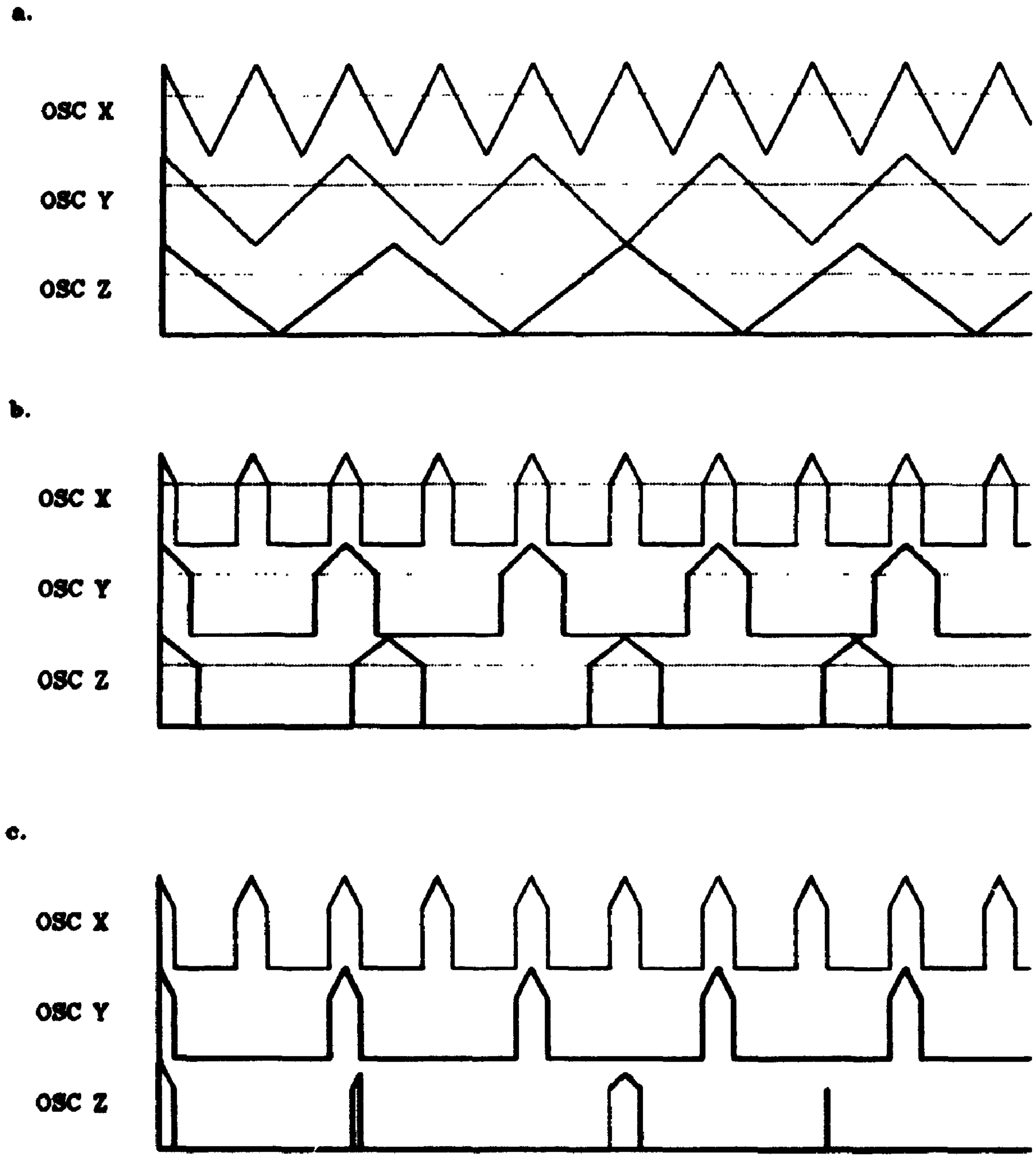


Figure 3.14. SQUARING and NEATENING operations. Figure 3.14a shows the ramp waveforms produced by three oscillators, X, Y and Z. Figure 3.14b shows these waveforms having been passed through SQUARING. Figure 3.14c shows the squared waveforms having been passed through NEATENING. For NEATENING, the fastest period functions as a gating trigger. In Figures 3.14a and 3.14b, the squaring index, Isq, is shown as a dotted line.

time is used to control which portions of the waveforms of other oscillators are passed through NEATENING. The squared outputs of all other oscillators are therefore switched off except when the energy level of the squared output of the fastest activated oscillator is high (not zero). The result of NEATENING on the squared set of waveforms produced in Figure 3.14b is shown in Figure 3.14c. As can be seen, the maxima of the squared waveforms of slower oscillators must coincide approximately with the maxima of the fastest period to pass through NEATENING. Since the maxima of each waveform represents the start of that oscillator's cycle, the periods of slower oscillators *must be in-phase integer multiples of the period of the fastest oscillator*. If a maximum of longer period is slightly out of synchrony with a maximum of the shortest period (i.e., the two periods are not exactly integrally related), only that portion of a longer squared waveform that coincides with the high portion of the fastest squared waveform will pass through NEATENING. Again this situation is illustrated in Figure 3.14.

Considering the operations of SQUARING and NEATENING together, it should also be noted that the width of the waveforms resulting from SQUARING depends on the squaring index, I_{sq}: the lower the value of I_{sq}, the wider the waveform. In turn, the width of the waveforms resulting from SQUARING will determine how precisely the maxima of longer periods must coincide with the trigger pulse to pass through NEATENING. Thus, the squaring index will determine how exactly the maxima of different periods must coincide to be categorized as integrally related. As mentioned earlier, for present purposes, the squaring index has been set at .6. However, this value may in fact vary from individual to individual, and from culture to culture reflecting varying personal and cultural sensitivities to mildly non-integrally related time intervals.

The output from the second sub-stage of the METER GATING (NEATENING) is a single channel carrying a complex waveform. This complex waveform is the aggregate of all the oscillators' waveforms that have passed through SQUARING and NEATENING. For example, the complex, aggregate waveform to result from the oscillator activity shown in Figure 3.14 is given in Figure 3.15.

One important general feature of aggregate waveforms should be noted: the more oscillators there are to entrain a particular period, the more prominent that period will be in the aggregate waveform. This observation leads in turn to three assertions that are of central importance to the model as a whole:

1. The most prominent or salient *individual* period in the aggregate waveform is chosen as the inter-beat time interval (the IBI).
2. The other time intervals are groupings and subdivisions of the beat.
3. The aggregate waveform as a whole *is* the listener's internal representation of the metrical structure of the musical stimulus.

Since the aggregate waveform may change over time depending on the stimulus, so the listener's representation of meter may also change over time. This changing representation is perceived real-time meter as defined in Chapter 1.

Returning to Figure 3.1, it can be seen that the output of METER GATING is fed to the OUTPUT BUFFER. This buffer is envisaged as working in similar fashion to SAMPLING BUFFER A in the SAS: sequentially separated events (now oscillator cycles) are integrated to form a

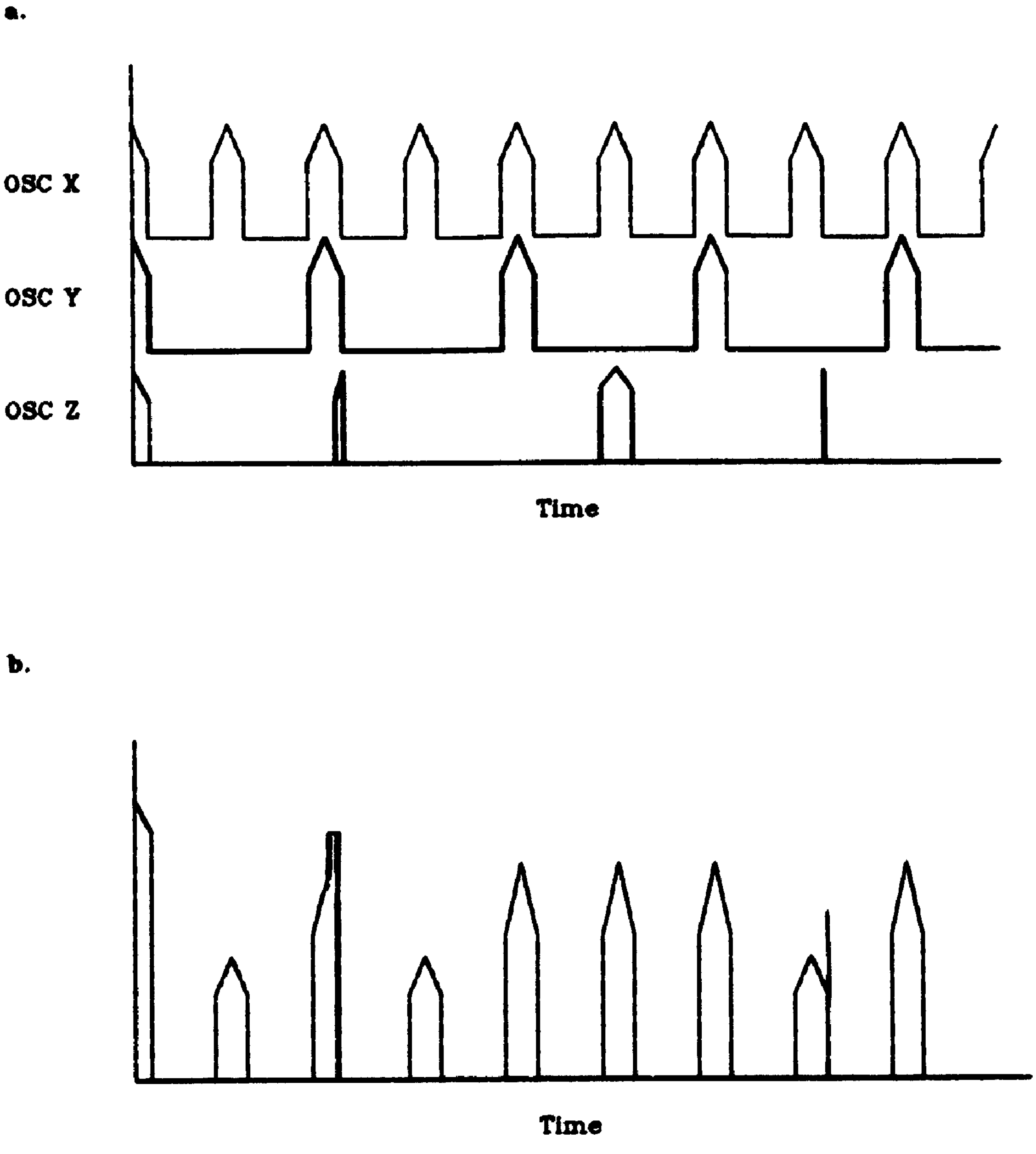


Figure 3.15. An aggregate waveform in the OUTPUT BUFFER. Figure 3.15a reproduces 3.14c, i.e., the output of NEATENING. Figure 3.15b shows the aggregate of these waveforms as they would be observed in the OUTPUT BUFFER. Note that a duple meter around the period produced by OSC X is evident, although this is disrupted somewhat by the intermittent presence of OSC Z's non-integer related period.

coherent "snap-shot" of the current representation of beat and meter. Like SAMPLING BUFFER A, the OUTPUT BUFFER has a variable storage capacity. In this model the OUTPUT BUFFER is ganged to SAMPLING BUFFER A, so that the storage capacity of the OUTPUT BUFFER depends on that of SAMPLING BUFFER A. In this way, the two buffers chunk their respective inputs in synchrony.

The output of the OUTPUT BUFFER is divided into two channels. As can be seen from Figures 3.1 and 3.11, the first channel is fed back to SAMPLING BUFFER B in the SAS. Thanks to this channel, the chunked representations of beat and meter that were integrated in the OUTPUT BUFFER are now stored with the raw acoustical data of their corresponding stimulus events in SAMPLING BUFFER B and other sampling buffers down the line in SAS. In this way, representations of beat and meter are stored in the same place as memory traces of raw acoustical data, and are subject to the same temporal compression in SAS that the signal traces themselves are. Thus, as noted in the earlier discussion of the SAS, the contents of the sampling buffers (from SAMPLING BUFFER B downwards) become a mixture of raw signal data and categorical information (abstract representations) about the signal.

The second output channel represents the output of the whole system and is fed to further processing. One possible feature of further processing might be an abstract long-term representation of meter (the *perceived abstract meter* defined in Chapter 1). Since the output of METER GATING is a constantly changing aggregate waveform, the contents of such long-term storage might be a set of idealized waveforms, against which the current contents of the OUTPUT BUFFER could be compared. These idealized waveforms would be "perfect" and unambiguous representations of meter gleaned from repeated and frequent exposure in the listener's lifetime. Two idealized waveforms for the meters 4/4 and 9/8 are shown in Figure 3.16. Thus, in labelling a particular stimulus as "4/4" or "9/8", the listener would match the pattern of his/her perceived real-time meter with a stored abstract representation of meter.

Literature Revisited

To conclude this chapter one can re-examine the list of specific issues arising from the literature review, and see how many of those may be accounted for by the model in some way. In some cases, these issues have already been addressed as part of the description of the model. In other cases, issues arising from the literature are addressed by synthesizing different aspects of the model for the first time.

a. The preferred beat tempo range between roughly 60 and 150 bpm (1000 and 400 ms). This range is accounted for by the output-period ranges of the oscillators in the OSCILLATOR BANK. Both the higher amplitudes and the closer bunching of oscillators around the value of 600 ms mean that periods between these approximate values will have a higher amplitude than periods outside this range. Recall that the oscillator period with the highest amplitude will be chosen as the beat level after it has passed through the METER GATING. It therefore follows, that if a variety of periods are entrained in the system described here, those in the 600 ms range will be chosen as the beat level.

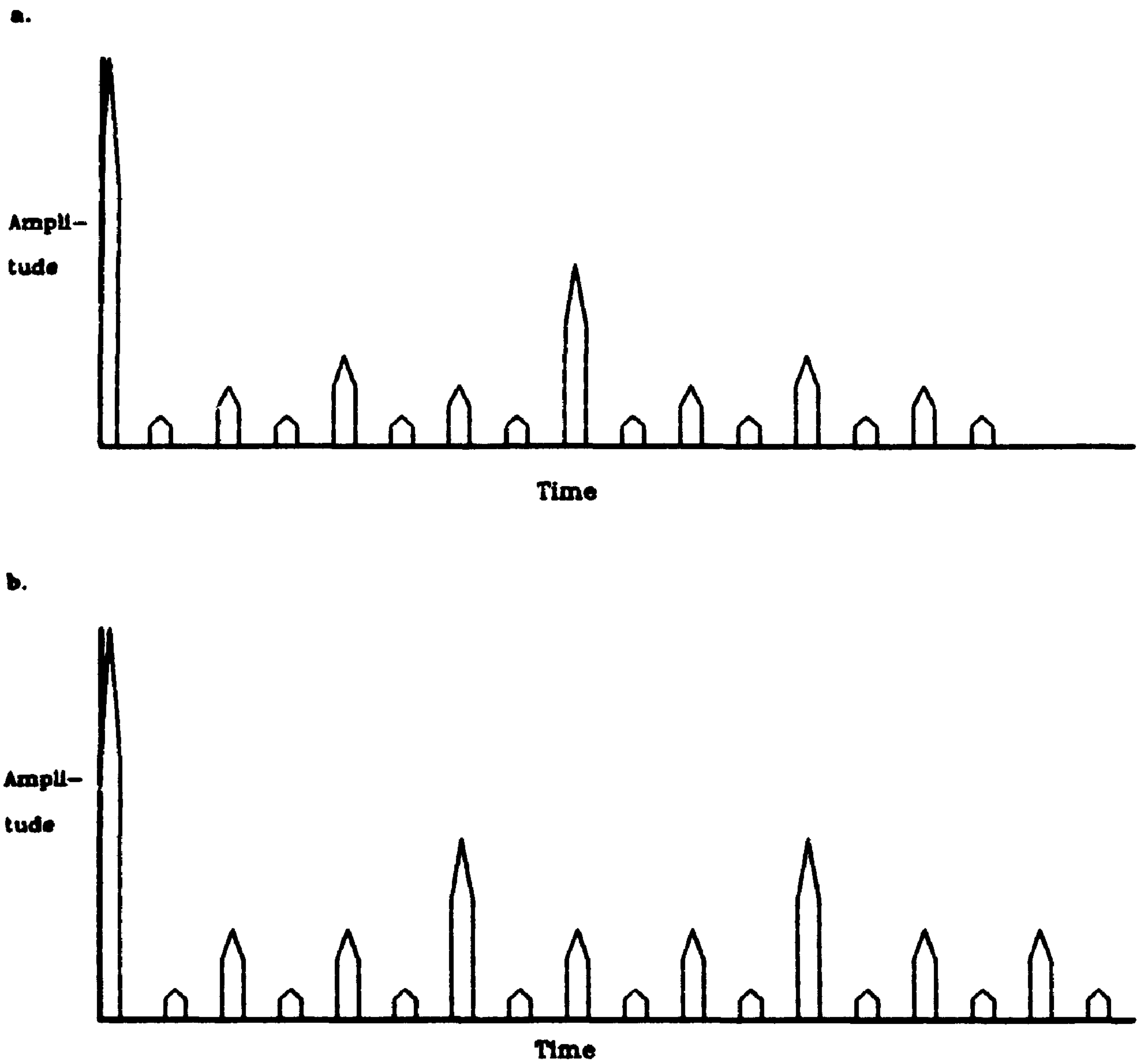


Figure 3.16. Idealized representations of two metrical structures. Figure 3.16a shows an idealized waveform for 4/4, from the sixteenth-note level up to the bar level. Figure 3.16b shows the same for 9/8. These waveforms are assumed to be abstracted from lifelong exposure to real-time representations of these metrical structures in the OUTPUT BUFFER.

b. Listeners' greater facility for hearing tempo decelerations over accelerations with simple, metronomic or computer-generated stimuli. When a tempo deceleration occurs over a number of beats, in effect the stimulus events are getting later and later with respect to the entrained periods in the hypothetical oscillators. However, if this change is gradual, the discrepancy between any two successive stimulus time intervals will be so small that the phase resetting that an oscillator must undergo has a very low probability of being noticed by the listener. For this reason, if a gradual change in tempo is noticed, it will not be in the OSCILLATOR BANK, but in the SAS by comparing the currently sounding beat rate (in SAMPLING BUFFER A) with a memory trace of the initial beat rate (in SAMPLING BUFFERS B, C, D etc.). As already mentioned, since memory traces in SAMPLING BUFFERS B, C, D etc. undergo temporal compression, gradual decelerations will be noticed more readily than accelerations of the same magnitude.

c. The interaction between metrical location of tempo change and direction of change as regards the ease of tempo change detection by listeners. This finding is not modelled by the proposed system.

d. The reasonably fine-grained detection of durational anomalies in perception (c. +/- 5%) as found by some experiments adopting the Weber's-law paradigm. Recall that in the model proposed here, time-interval anomalies will result in phase resetting of an entrained oscillator. The overall probability of this resetting being noticed by the listener is determined by the value of OP_R, as previously defined. As with probability values in null-hypothesis rejection/acceptance in experimental statistics, one can determine a cut-off level (alpha level) beyond which an occurrence is deemed unlikely to arise. With OP_R one can also decide on an arbitrary threshold level below which a listener is deemed unlikely to notice a rhythmic discrepancy. The threshold for P_R can be set at any level desired, thus making the system's detection of rhythmic anomalies as fine-grained as one wishes. This variability could reflect the variability of rhythmic anomaly detection that might arise amongst individuals owing to musical training or ability.

e. The much less fine-grained reliance on integrally related IOIs in rhythm cognition. Rhythm cognition, as opposed to rhythm perception, will only take place once specific information has been abstracted from the raw stimulus data. The output aggregate waveform from the METER GATING is just such an abstraction. The stipulation in the NEATENING sub-stage that trimmed oscillator waves are only passed through if they coincide with the trigger (shortest period), guarantees that all components of the aggregate waveform are integer multiples of the shortest period. It may well be the case that it is this aggregate waveform that controls motor responses to rhythmic stimuli as well as being the part of the system most accessible to conscious introspection in general. Since, in experimentation on rhythm cognition, the types of data gathered are generally either motoric responses or verbal labels ("same/different," "3/4 versus 6/8"), it is easy to see how a reliance on integer related time intervals would be commonplace.

f. The tendency by musically more sophisticated listeners to parse a beat from an input stimulus to provide a reference time interval around which all other time intervals are organized.

g. The rapid establishment of a metrical hierarchy as multiples and divisors of the beat time interval.

These two issues may be considered together. The establishment of a beat and a metrical structure is precisely what the hypothetical system does. Entrainment allows the beat time interval to be established after only two or three stimulus events have occurred. As soon as more than one metrical level is present in the stimulus, the system responds by entraining a second period, resulting in some type of metrical representation in the form of a complex waveform on the output channel of the METER GATING. The fact that musically trained listeners (at least Western listeners) find it easier to establish a metrical hierarchy than musically untrained listeners may be due to two factors: (i) their IAIT is better able to extract accent cues from the SAS, and (ii) their SQUARING sub-stage has a lower squaring threshold, thus making their METER GATING less stringent about integer-related periods.

h. The tendency to categorize time intervals not only as integrally related, but also to categorize them as conforming to an established metrical context. This issue is a more complex version of (f) above. In order to be passed through METER GATING, oscillator periods must be (approximate) integer multiples of the fastest oscillator period. This stipulation will ensure that the output of METER GATING conforms to an established meter, so long as the fastest oscillator period is itself part of the established metrical framework. In both of the Clarke (1987a) stimuli the fastest stimulus time interval *is* part of the metrical framework: 320 ms equals an eighth-note in the 6/8 context, and 240 ms equals a sixteenth-note in the 2/4 context. So far as the present author is aware, there are no empirical data concerning whether categorical perception will still occur when a metrically *non*-conformant time interval is also the fastest time interval in a stimulus, and therefore the trigger for NEATENING. The present model predicts that categorical perception would be disrupted in such a situation.

Conclusion

The general goals outlined at the beginning of this chapter were: (1) to account for as much of the literature as possible, (2) to generate specific experimental hypotheses, (3) to provide a clear conceptualization of rhythm perception.

Goals one and three may be considered together. Firstly, the degree to which this model accounts for the findings of the literature reviewed may be judged from the previous paragraphs. Doubtless some sources of variance are not accounted for, however the model does seem to fall in line with major points of convergence in the literature. As regards goal three, only the reader can judge if this model provides a useful way of thinking about beat and meter perception.

One can close by addressing goal two, that is, by summarizing several specific hypotheses arising from the interaction of stages in this model:

a. Large discrepancies in IOIs, that is, sudden changes in time interval are noticed in the OSCILLATOR BANK.

b. Subtle discrepancies in IOIs, that is, gradual changes, if they are noticed at all, are detected by comparing perception and memory (the contents of the SAS).

c. With simple stimuli, at least, tempo decelerations will be noticed more easily than accelerations of the same magnitude.

d. Integer-related time intervals will be handled more easily by the cognitive system than non-integer related time intervals.

e. The OSCILLATOR BANK will handle time intervals in the middle of its period-response range (i.e., around 600 ms) better than time intervals at the extremes of its range.

Some of these hypotheses will be expanded, refined and tested in the next chapter, Chapter 4, which introduces three related experimental studies.

CHAPTER 4

EXPERIMENTAL DESIGN AND METHOD

This chapter is concerned with the design, the preparation, and the running of the three experiments that form the empirical part of this dissertation. The first part of this chapter deals with the experimental hypotheses and the theoretical considerations pertaining to the testing of these hypotheses. The second part treats the practical aspects of preparing and performing the three experiments, and includes the type of detailed information required for replication.

Part I - Design

Experimental Hypotheses

In order to introduce the experimental design employed in this study, it is advantageous to reconsider the principal hypothesis of this dissertation, as it was formulated in Chapter 1:

- Temporal changes that occur suddenly, i.e., within the span of the psychological present, are detected in perception; temporal changes that occur gradually, i.e., outside the span of the psychological present are detected by comparing a memory trace with perception.

From the definitions given in Chapter 1, rhythm modulation would constitute a sudden change occurring within the psychological present, whilst tempo modulation would constitute a gradual change that could only be detected in memory.

The model described in Chapter 3 throws more light on what specific perceptual/cognitive mechanisms might be meant by "perception" and "memory." According to the model, temporal changes detected in perception are detected because they cause a phase resetting of one or more of the oscillators in the OSCILLATOR BANK. Temporal changes detected in memory are detected because a trace of a previous stimulus stored in one or more of the later sampling buffers of the SHORT AUDITORY STORE (SAS) differs markedly from the stimulus currently held in SAMPLING BUFFER A. The hypothetical workings of the OSCILLATOR BANK and the SHORT AUDITORY STORE, designed to account for the findings of the literature reviewed in Chapter 2, also give rise to the four hypotheses that are tested in the experiments described in this chapter. The four experimental hypotheses are outlined in the following paragraphs, along with the reasoning behind them.

Hypothesis 1. For rhythm modulation, a late anomalous onset and an early anomalous onset will be equally difficult to detect. However, for tempo modulation, an early anomalous onset will be more difficult to detect than a late anomalous onset.

The first part of Hypothesis 1 arises from the workings of the OSCILLATOR BANK. As defined in Chapter 3, the probability of detecting the resetting of an oscillator, P_r, is a function of the instantaneous energy level, EL_i, of the oscillator's waveform at the instant before phase resetting. Since the ramp waveform of each oscillator is envisaged as symmetrical, a resetting caused by an early anomalous stimulus onset and one caused by a late anomalous stimulus onset of the same magnitude will occur at identical EL values on the oscillator's output waveform. This situation is illustrated in Figure 4.1. For this reason, early and late anomalous stimulus onsets of the same magnitude are equally likely to be noticed by the listener.

The second part of Hypothesis 1 stems from the workings of the SAS. Since traces in the later sampling buffers of the SAS are compressed temporally, the time interval (or the representation of time interval) separating two events would be shorter in these later stages of memory than in perception. Therefore the representations of stimuli in later stages of memory will seem to have accelerated relative to representations of the same stimulus in perception and earlier stages of memory. Tempo accelerations will therefore not be detected so easily as decelerations of the same magnitude since the compressed memory trace will be more similar to an accelerated beat rate in perception than to a decelerated one. In short, the difference between memory and perception will be less when a beat has accelerated compared to when a beat has decelerated.

Hypothesis 2. Rhythm modulation at the extremes of the beat-rate range will be more difficult to detect than modulation in the middle of this range. Tempo modulation at slow beat rates will be more difficult to detect than tempo modulation at fast beat rates.

The first part of Hypothesis 2 springs from the fact that mid-range tempi fall most readily into the preferred tempo range identified in Chapter 2. Beats in the middle of the tempo range, i.e. around 600 ms (100 bpm), will, on average, entrain the greatest number of

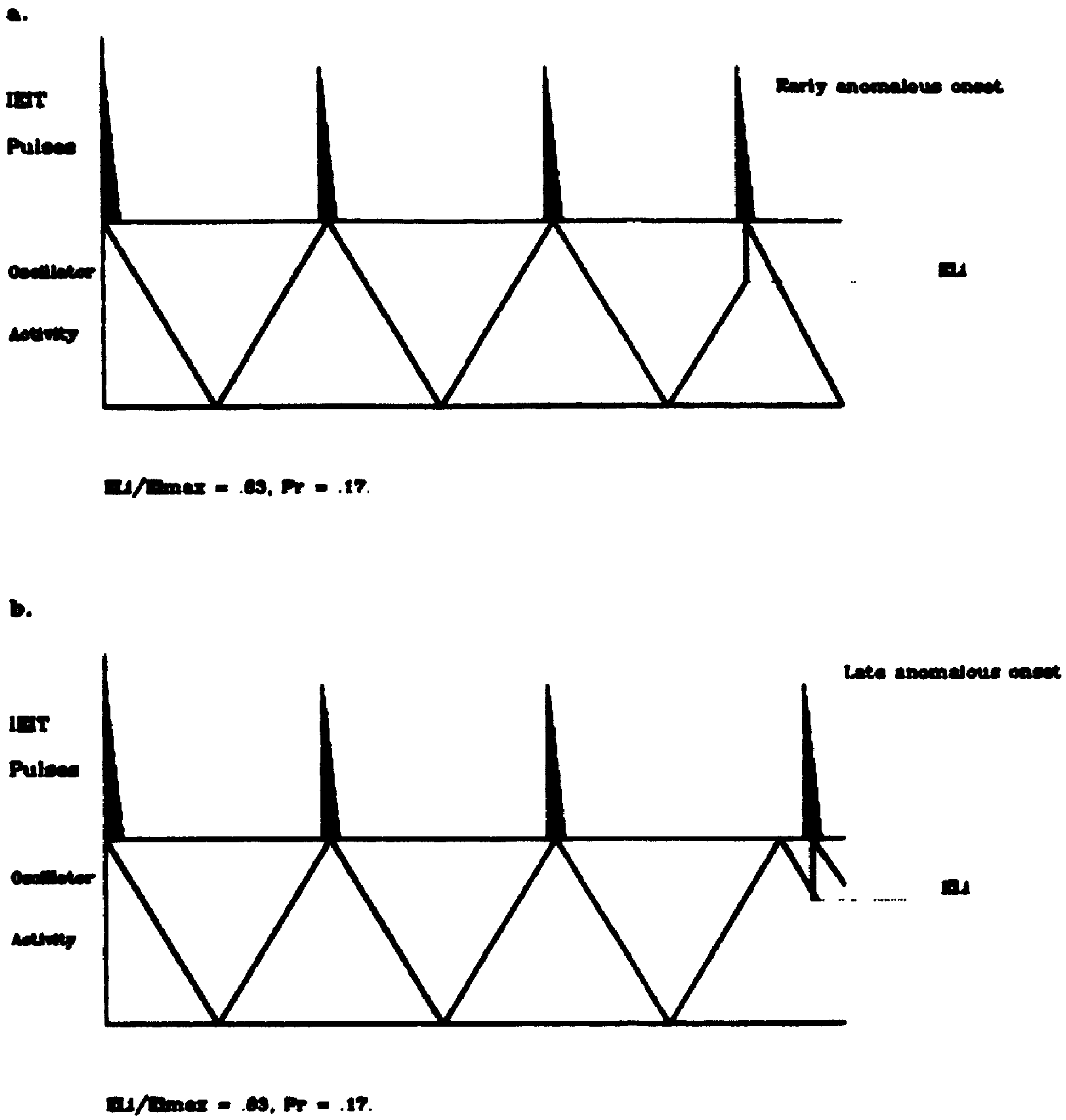


Figure 4.1. Phase resetting caused by early and late anomalous onsets. Figure 4.1a shows resetting with an early anomalous onset, Figure 4.1b shows the same with a late anomalous onset. Note that since the asynchrony of the anomalous onsets relative to the already-entrained period is the same in both cases, the unweighted probability value, Pr, is identical for the two situations.

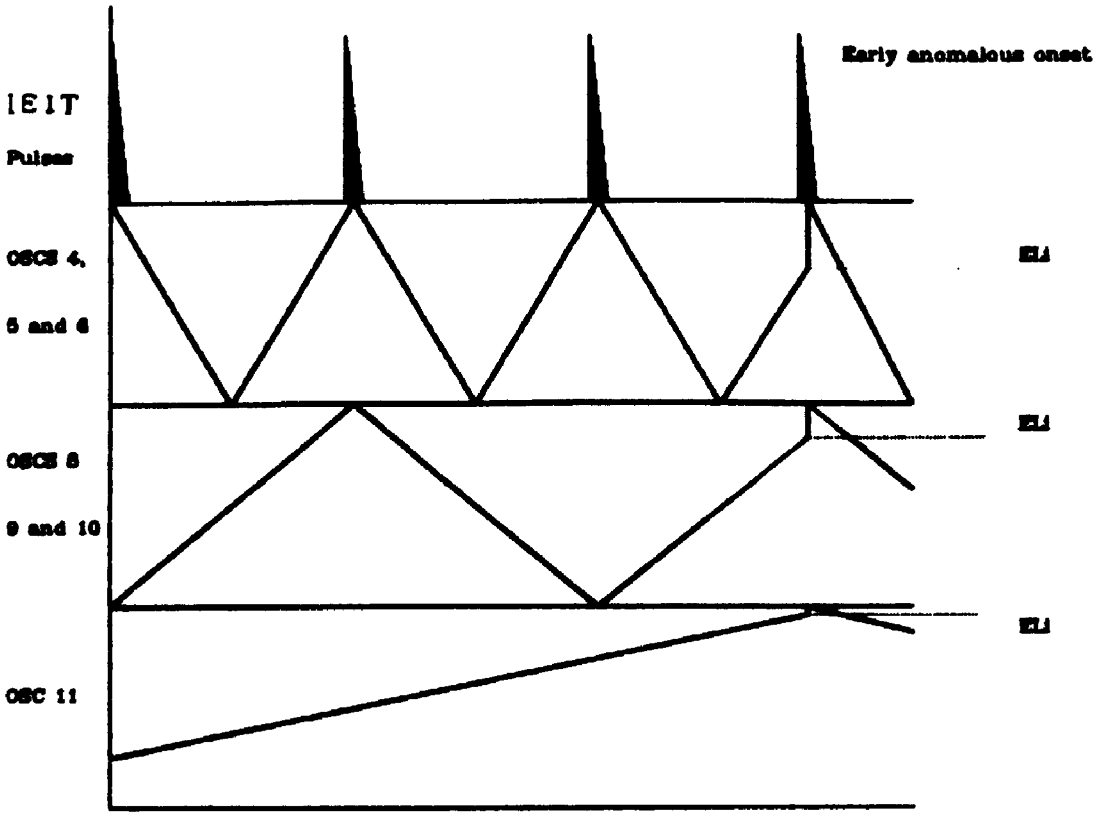
oscillators at the highest amplitudes. If several oscillators are entrained to a single stimulus time interval, when an anomalous time interval is presented to the listener, all these oscillators will phase reset together. The greater the number of oscillators oscillating in sychrony the higher will be the overall probability (OP_R) of a phase resetting being detected.

The second part of Hypothesis 2 derives primarily from the literature (particularly Ellis 1991 and Royal 1993 [see Appendix A]), which has observed a greater facility in detecting tempo modulation at faster tempi. In terms of the current model, one possible explanation for this behaviour may be as follows: the faster the stimulus, the greater the number of beats that can be stored as a single chunk in one sampling buffer in the SAS. The greater the number of beats to be stored in one chunk, the less the system will have to pass signal traces "on the line" to later sampling buffers. The less the reliance on later sampling buffers, the less compressed and distorted is the memory trace of the beginning of the stimulus. The less distorted the memory trace of the initial beat rate, the more accurate is the process of comparison and detection of tempo modulation.

Of course, contrary to the above, it may be argued that faster stimuli constitute more complex stimuli, and, therefore, with faster tempi, SAMPLING BUFFER A will sample at a correspondingly faster rate, reducing its overall storage capacity. If SAMPLING BUFFER A has a reduced storage capacity, the quicker it would pass stimulus chunks down the line, cancelling out any possibility that more stimulus events could be encoded in one chunk. However, with the extremely simple stimuli that have so far been used in the experiments reported (isochronous, monophonic melodies, metronome ticks or drum strikes), even at a fast tempo, there may be no real need for SAMPLING BUFFER A to increase its sampling rate by any substantial amount in order to encode a signal trace of sufficient fidelity. Therefore more events could still be encoded in each chunk in SAS.

Hypothesis 3. For rhythm modulation, an anomalous onset on beat 1 of the bar will be more difficult to detect than the same on beat 2. For tempo modulation, there will be no difference between an anomalous onset on beat 1 and beat 2.

The reason for the first part of Hypothesis 3 is as follows: if a stimulus implies a metrical hierarchy, then different oscillators will be entrained to the different time intervals represented by this hierarchy. For example, with the current model, if OSCILLATORS 4, 5 and 6 entrain to a quarter note of 500 ms time interval in a 4/4 meter, OSCILLATORS 8, 9 and 10 can entrain to the half-note time interval (1000 ms), and OSCILLATOR 11 can entrain to the whole-note time interval (2000 ms). Therefore, beat 1 in the bar will witness the maxima of not only OSCILLATORS 4, 5 and 6 (the beat level) but also those of OSCILLATORS 8, 9, 10 and 11 too. On the other hand, the only maxima at beat 2 in the bar will be those of OSCILLATORS 4, 5 and 6. As can be seen from Figure 4.2, an anomalous onset that causes substantial resetting of oscillators 4, 5 and 6, will, however, cause only a negligible resetting of oscillators oscillating at longer periods. As noted in Chapter 3, when several groups of oscillators are oscillating at contrasting periods, the listener may attend to any one of these. If the listener is



OSCS 4,5 and 6: ELi = .83, Pr = .17

OSCS 8, 9 and 10: ELi = .9; Pr = .1

OSC 11: ELi = .98; Pr = .02

Figure 4.2. Different Pr values registered by different oscillators for the same anomalous onset. Since oscillators of contrasting periods produce output waveforms of contrasting gradients, the Pr value at a given point of reset will be smaller for those oscillators oscillating at longer periods. Note that the Pr value is the unweighted probability for each individual oscillator.

attending to the outputs of the slower oscillators when the anomalous onset occurs, he/she will be less likely to notice the anomaly. On beat 2, the listener's attention cannot be divided in the same way. Therefore, in stimuli that imply several metrical levels longer than the beat, anomalous onsets on beat 1 will, on average, be more difficult to hear than anomalous onsets of the same magnitude on beat 2.

The reason for the second part of Hypothesis 3 is simply that, since, as previously argued, tempo modulation will not be detected in the OSCILLATOR BANK, it seems unlikely that the metrical location of the start of a tempo modulation could affect the trace in the SAS. In the current model, both raw acoustical signal data and representations of beat and meter are stored in the SAS from SAMPLING BUFFER B downwards. However, the extent to which the contents of a sampling buffer are temporally compressed depends on how far down the line the sampling buffer is (i.e., how old are its contents). This process is carried out on all stimulus time intervals, regardless of their metrical position.

Hypothesis 4. For rhythm modulation, anomalous onsets in stimuli with no subdivision of the basic beat will be more difficult to detect than stimuli where the beat IS subdivided. For tempo modulation, it is not known whether presence or absence of beat subdivision will effect detection of anomalous onsets.

The reason for the increased facility in detecting rhythm modulation with a subdivision is essentially an extension of the argument for Hypothesis 3. For a subdivided stimulus, oscillators will entrain to the beat level, and the subdivision level. For example, with the current model, if OSCILLATORS 4, 5 and 6 entrain to a quarter note of 500 ms time interval in a 4/4 meter. OSCILLATORS 1 and 2 can entrain to the eighth-note time interval (250 ms). As illustrated in Figure 4.3, an IOI anomaly that is only just noticeable at the beat level (500 ms), will be very apparent at the subdivision level (250 ms).

In the case of tempo modulation, it is impossible to predict the effect (if any) of beat subdivision on modulation detection since either one of two scenarios is feasible. On the one hand, the increased complexity of the subdivided stimulus could bring about a faster sampling rate and therefore shorter chunks in each sampling buffer. The upshot of this process would be a greater reliance on buffers "further down the line" and therefore the accuracy of comparisons would be generally impaired. On the other hand, if the sampling rate does not increase greatly owing to stimulus subdivision, then a subdivided stimulus will mean more temporal information in each sampling buffer. The more information in each buffer, the more temporal points of reference the listener would have for the purposes of comparison between old and new traces. The greater the number the points of comparison the more accurate would be the detection of tempo modulation.

Stimuli and Independent Variables

In order to test the above hypotheses, Experiments 1, 2 and 3 examined the perception of rhythm and tempo modulation under different conditions. Experiment 1 was designed to test Hypotheses 1 and 2. Experiment 2 was designed to test Hypotheses 1, 2 and 3. Experiment 3 was designed to test Hypotheses 1 and 4. The stimuli and independent

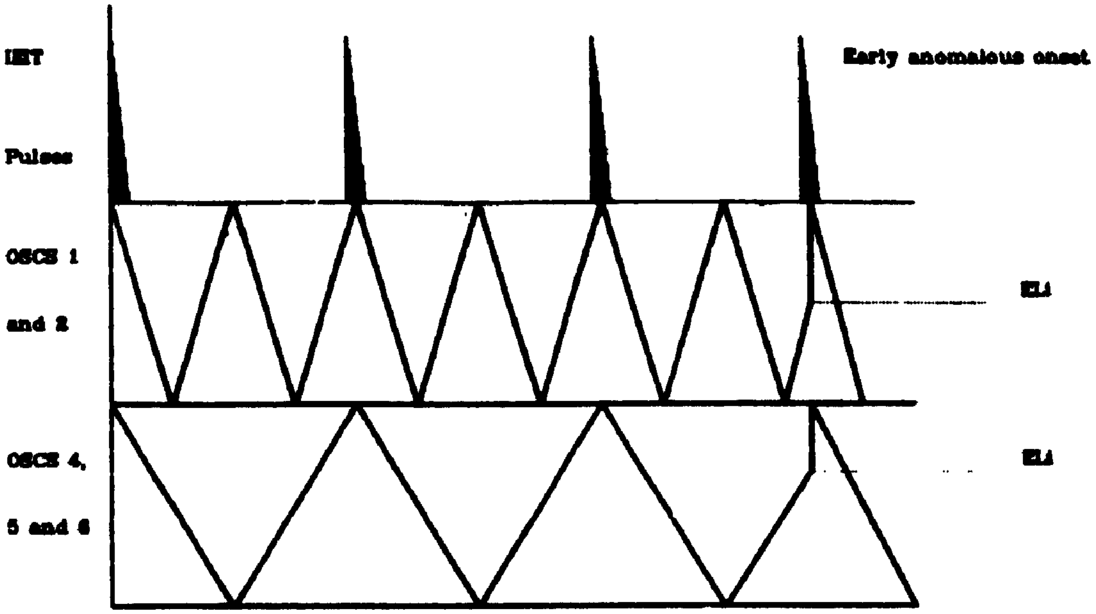


Figure 4.3. Phase resetting with a subdivided beat. As in Figure 4.2, owing to the different gradients of the output waveforms of oscillators of contrasting periods, an oscillator that has entrained a beat subdivision will register a larger Pr value than those oscillators that are entrained to the beat time interval.

variables that were used for each experiment are outlined in the following paragraphs.

Experiment 1

General Trial Format. Each trial consisted of a series of 32 beats (a synthesized unpitched drum sound), with a heavy dynamic accent added to the first of every four beats, and a medium dynamic accent added to the third of every four beats. The result of this combination of beats and accents was 8 bars of quadruple meter, with three levels of metrical hierarchy: the time interval *i* between successive beats, the time interval *2i* between successive accents, and the time interval *4i* between successive heavy accents (downbeats). In each trial, at least the first bar was sounded isochronously without any rhythm or tempo modulation. However, on the downbeat of either bar two or bar three (determined by random selection), a gradual modulation in either rhythm OR tempo was introduced (or the rhythm remained isochronous at a constant tempo). The modulation continued for six bars after which the stimulus remained constant at the new rhythm/tempo until the end of the trial. Figure 4.4 shows the general format for both stimuli with modulation beginning in bar two as well as bar three. Trials were manipulated according to the following independent variables:

TYPE. TYPE had two levels, *rhythm* modulation and *tempo* modulation. For *rhythm* modulation, the inter-onset time intervals (IOIs) between the first and the second beats, and between the second and third beats of each bar became increasingly more uneven. The IOIs between the third and fourth, and fourth and first beats remained

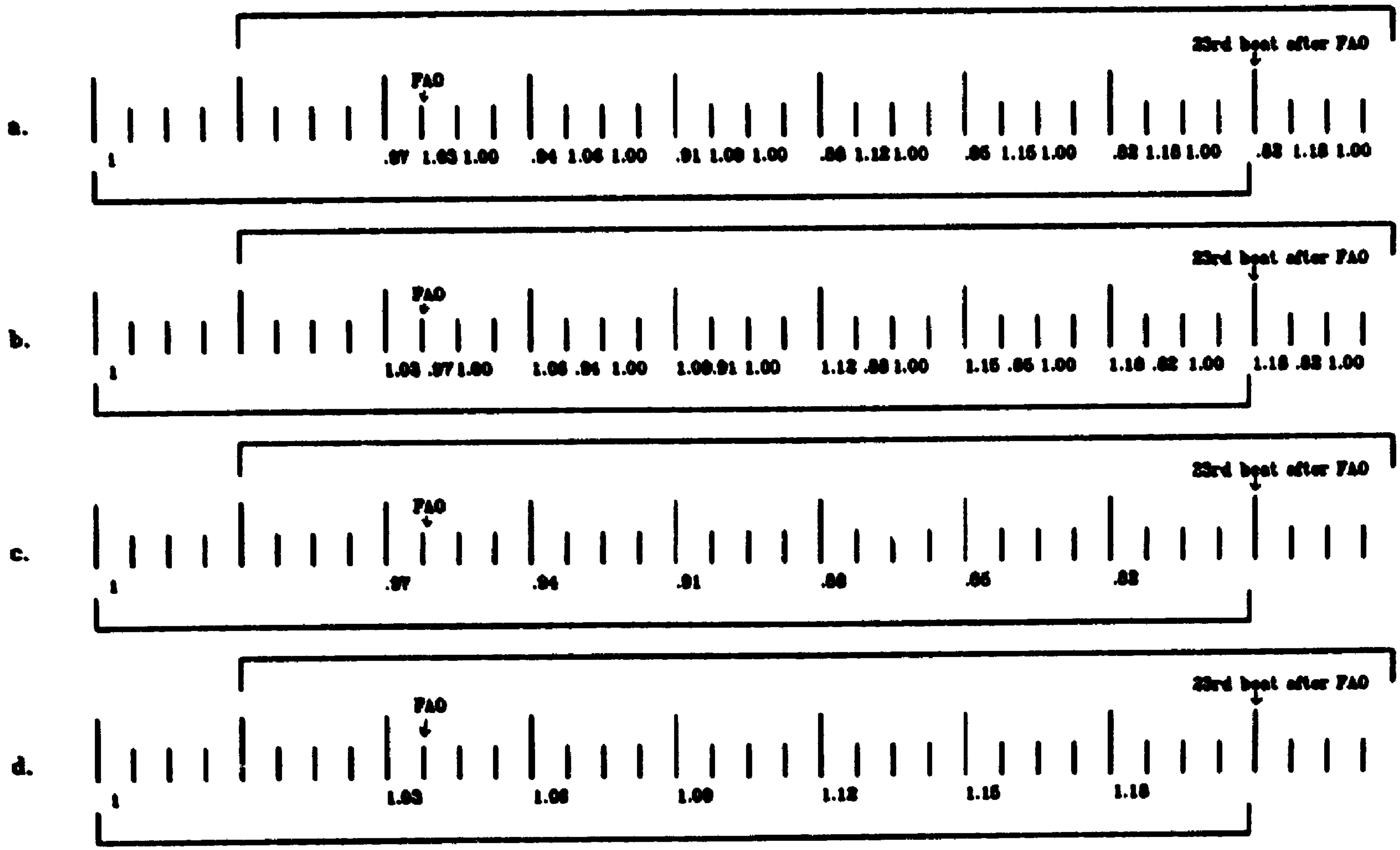


Figure 4.4. Generic stimuli for Experiment 1. Figure 4.4a shows a combination of rhythm modulation with an early FAO. Figure 4.4b shows a combination of rhythm modulation with a late FAO. Figure 4.4c shows a combination of tempo modulation with an early FAO. Figure 4.4d shows a combination of tempo modulation with a late FAO. For all stimuli, the square brackets above the event onsets represents the span of the complete stimulus with the FAO in bar 2; conversely, the square bracket below the event onsets represents the span of the complete stimulus with the FAO in bar 3. Numbers below the event onsets are proportions of the initial IOI, i.

constant at the initial beat rate. Examples of *rhythm* modulation are given in Figures 4.4a and 4.4b. For *tempo* modulation, the beat rate changed at the beginning of each bar and was maintained at the new rate for the time interval of one bar. Overall, this procedure resulted in a gradual change of tempo away from the initial beat rate. Examples of *tempo* modulation are given in Figures 4.4c and 4.4d.

DIRECTION. DIRECTION had two levels, *early* and *late*, depending on when the first anomalous onset in each bar occurred relative to preceding beats. For the *early* condition, the IOI preceding the anomalous onset was shortened by decrements of 3% of the initial IOI per bar. Thus, if i was the initial IOI then successive shortened IOIs would be .97i, .94i, .91i, .88i, .85i and .82i. For the *late* condition, the IOI preceding the anomalous onset was lengthened by increments of 3% of the initial IOI per bar, that is 1.03i, 1.06i, 1.09i, 1.12i, 1.15i and 1.18i.

The combination of the *rhythm* and *early* conditions, resulted in a rhythmic stimulus that modulated from an isochronous sequence to one where the first two beats of each bar became iambic. The combination of *tempo* and *early* conditions resulted in a rhythmic stimulus where the beat rate accelerated gradually. Examples of early anomalous onsets are given in Figures 4.4a and 4.4c. The combination of *rhythm* and *late* conditions resulted in a rhythm that became increasingly more trochaic. The combination of *tempo* and *late* conditions resulted in a stimulus whose tempo decelerated gradually. Examples of late anomalous onsets are given in Figure 4.4b and 4.4d.

INITIAL BEAT RATE. INITIAL BEAT RATE had five levels, differentiated by the IOIs between the isochronous drum-beats at the very beginning of each trial. This IOI is designated as "i" in Figure 4.4. Since, of course, beat rate is inversely proportional to time interval, the longer was i the slower was the INITIAL BEAT RATE. The five levels of INITIAL BEAT RATE were: *45* bpm (i = 1333.3 ms), *75* bpm (i = 800 ms), *120* bpm (i = 500 ms), *190* bpm (i = 315.8 ms) and *300* bpm (i = 200 ms). These values were intended to sample the entire musical tempo range.

VERSION. VERSION had two levels, *original* and *replication*. The inclusion of VERSION as a variable provided two renditions of each condition. The purpose of this was to allow the calculation of a measure of reliability for one subject over trials. If the same trends with respect to the variables TYPE, DIRECTION and INITIAL BEAT RATE were reliably observed across two renditions of the same stimulus, then the experiment would be a reliable measure of difficulty of temporal change detection.

In addition, the VERSION variable allowed the experimenter to monitor any possible changes in subject performance across trials, owing to extraneous, uncontrolled factors. If, on average, a subject's performance deteriorated between two renditions of the same condition, then this might be attributed to fatigue. Secondly, if his/her performance evinced a general improvement between the original and replication versions, then this might be evidence for learning. Thirdly, if no significant difference in performance occurred between original and replication version, then one of two conclusions is possible: (1) neither fatigue nor learning occurred; (2) both fatigue and learning occurred together, cancelling each other out. No hypothesis could be ventured in advance concerning which scenario might occur.

The above independent variables resulted in a 2 (TYPE) by 2 (DIRECTION) by 5 (INITIAL BEAT RATE) by 2 (VERSION) factorial design, giving 40 trials. In addition, a constant trial, containing no rhythm or tempo modulation, was included at each of the five initial beat rates, giving 45 trials per subject. Constant trials were included to give subjects the possibility of a third choice when it came to classifying the type of modulation that occurred. Since all three experiments were concerned with the detection of temporal change, data from the constant trials were disregarded in the statistical analysis.

Experiment 1 was designed to test Hypotheses 1 and 2. These hypotheses may now be rephrased in terms of the independent variables listed above:

- Hypothesis 1. There will be a statistically significant interaction between TYPE and DIRECTION. *Rhythm* modulation will be indifferent to direction of modulation (*early* versus *late*); *tempo* modulation will be more difficult with *early* than *late*.
- Hypothesis 2. There will be a statistically significant interaction between TYPE and INITIAL BEAT RATE. *Rhythm* modulation will be most difficult to detect at extreme initial beat rates, whereas *tempo* modulation will be most difficult at slow initial beat rates.

If Hypotheses 1 and 2 are confirmed, these results would be consistent with the notion that rhythm and tempo modulation are processed by different perceptual mechanisms, and that these mechanisms manifest different measurable behaviours.

Experiment 2

General Trial Format. As in Experiment 1, each trial consisted of a series of 32 drum beats with dynamic accents added to produce an unambiguous 4/4 meter. However, this time, in each trial at least the first three beats were sounded isochronously without any rhythm or tempo modulation. If modulation occurred, the first anomalous onset then occurred in bar two or bar three (randomly selected) on either the first or the second beats of these bars. As in Experiment 1, the modulation continued for six bars after which the stimulus remained constant at the new rhythm/tempo until the end of the trial. Figure 4.5 shows the general format for both stimuli with modulation beginning on beat 1 in bar two as well as bar three. Beat-2 stimuli were essentially the same as those stimuli used for Experiment 1, as illustrated in Figure 4.4. Trials were manipulated according to the following independent variables:

TYPE. As in Experiment 1, TYPE had two levels, *rhythm* modulation and *tempo* modulation.

DIRECTION. As in Experiment 1, DIRECTION had two levels, *early* and *late*, depending on when the first anomalous onset in each bar occurred relative to preceding beats. Also as in Experiment 1, the *early* condition involved IOIs getting shorter by 3% of the initial IOI per bar, and the *late* condition involved IOIs getting longer by 3% per bar.

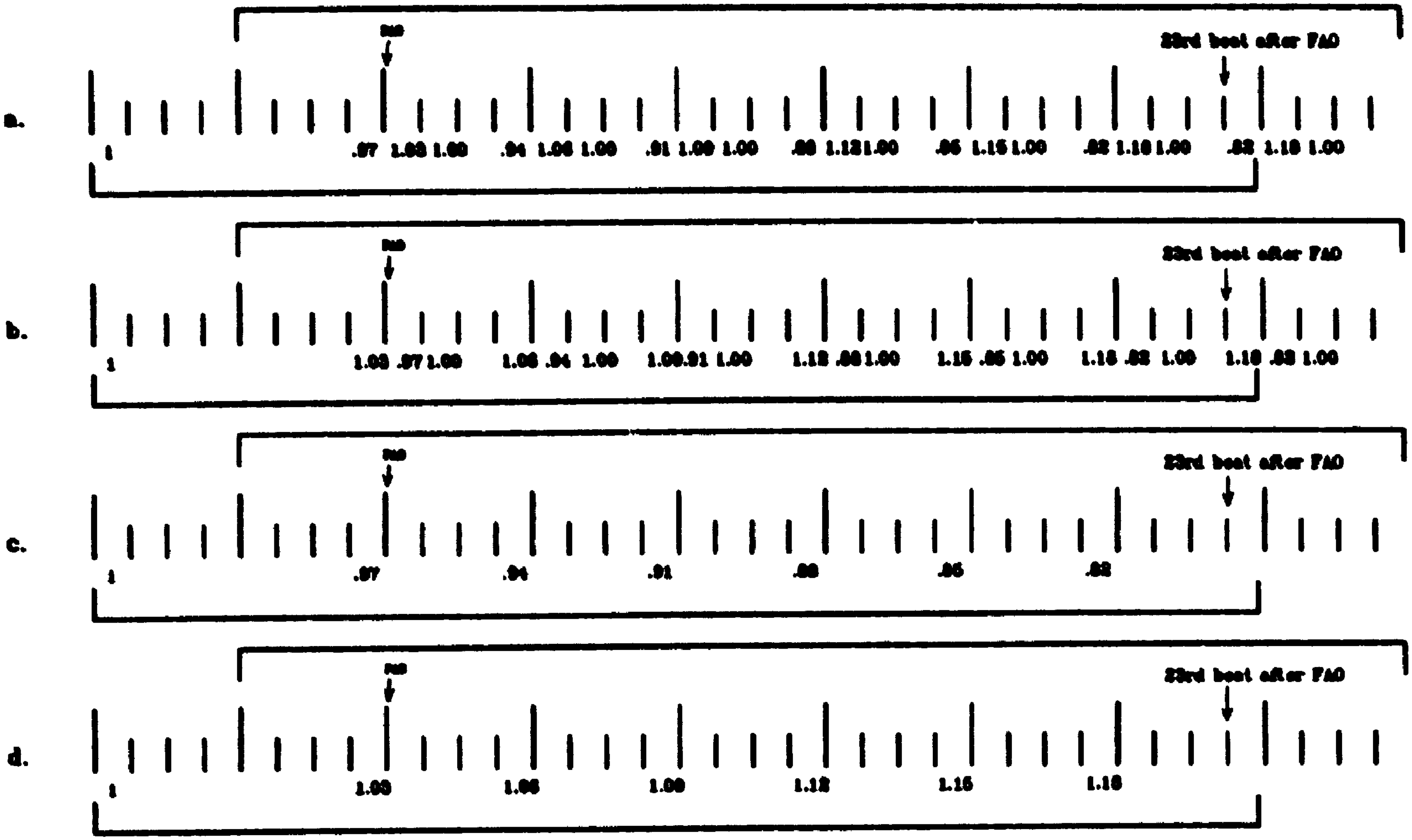


Figure 4.5. Generic stimuli for Experiment 2, with FAO on beat 1 of the bar. Figure 4.5a shows a combination of rhythm change with an early FAO. Figure 4.5b shows a combination of rhythm change with a late FAO. Figure 4.5c shows a combination of tempo change with an early FAO. Figure 4.5d shows a combination of tempo change with a late FAO. For all stimuli, the square bracket above the event onsets represents the span of the stimulus with the FAO in bar 2; conversely, the square bracket below the event onsets represents the span of the stimulus with the FAO in bar 3. Proportional IOIs are shown as in Figure 4.4.

INITIAL BEAT RATE. In Experiment 2 INITIAL BEAT RATE had only two levels: *60* bpm (i = 1000 ms), and *150* bpm (i = 400 ms). These values sample a narrower range than those used in Experiment 1. However, they represent the approximate outer limits of the preferred tempo range as identified in Chapter 2.

LOCATION. LOCATION represents the most important difference between Experiments 1 and 2. In Experiment 2, LOCATION had two levels, *beat 1* and *beat 2*. In the *beat-1* condition, the first anomalous onset in each bar occurred on beat 1, the downbeat. In the *beat-2* condition, the first anomalous onset in each bar occurred on beat 2. Although the INITIAL BEAT RATE (and therefore the absolute value of i) is different, the *beat-2* conditions in Experiment 2 are essentially the same as the stimuli used in Experiment 1 and illustrated in Figures 4.4. The beat 1 conditions for Experiment 2 are illustrated in Figure 4.5.

VERSION. As in Experiment 1, VERSION had two levels, *original* and *replication*. Again, the inclusion of two renditions of each condition allowed the calculation of a measure of reliability and the identification of either fatigue or learning as an extraneous factor in the experiment.

The independent variables in Experiment 2 resulted in a 2 (TYPE) by 2 (DIRECTION) by 2 (INITIAL BEAT RATE) by 2 (LOCATION) by 2 (VERSION) factorial design, giving 32 trials. In addition, two constant trials were included for each of the two initial beat rates, giving 36 trials per subject. Again, constant trials were included merely as a foil, and data from them were disregarded in the statistical analysis.

merely as a foil, and data from them were disregarded in the statistical analysis.

Experiment 3 was designed to test Hypotheses 1 and 4, which may be re-expressed in terms of the above independent variables:

- Hypothesis 1. There will be a statistically significant interaction between TYPE and DIRECTION. *Rhythm* modulation will be indifferent to direction of modulation (*early* versus *late*); *tempo* modulation will be more difficult with *early* than *late*.
- Hypothesis 4. There will be a statistically significant effect owing to PIECE in the *rhythm* conditions. Owing to the lack of beat subdivision, *rhythm* modulation in the *Author* piece will be more difficult to detect than equivalent modulation in the three pieces with beat subdivision. It is unknown how PIECE will affect *tempo* modulation perception.

Dependent Variables

The purpose of all three experiments was to measure the relative difficulty of detecting rhythm and tempo modulation in the various contexts prescribed by the independent variables outlined above. In all three experiments, two dependent variables were used that reflected difficulty in detecting an IOI anomaly: (1) beats-to-respond data and (2) right/wrong answer data. Each dependent variable is considered in turn.

"Prelude No.5 in D major" from Book 1 of *The Well-Tempered Clavier* by Bach, "Étude Op.10, No.2 in A minor" by Chopin, and "Interludium [No.4]" from *Ludus Tonalis* by Hindemith. The specially-composed piece was written in a common-practice tonal style, its only unusual feature being that it maintained an isochronous quarter-note rhythm in its upper voice throughout. The incipits of three of the four pieces that were used are given in Figure 4.6.

Each trial consisted of the first eight bars of one of the above pieces, played on a synthesized piano sound with minimal dynamic and articulatory accents in the Bach, Chopin and Hindemith pieces, and similar dynamic accents to the stimuli of Experiments 1 and 2 in the specially composed piece. In each trial, at least one complete bar of 4/4 (16 sixteenths) was sounded isochronously without any modulation of rhythm or tempo. However, on the downbeat of bar 2 or 3, determined at random, modulation in either tempo or rhythm was introduced (or the rhythm remained constant). The modulation continued for six bars, and then the stimulus remained constant at the new rhythm/tempo until the end of the extract. The trials in Experiment 3 were manipulated according to the following variables:

TYPE. As in Experiments 1 and 2, TYPE had two levels, *rhythm* modulation and *tempo* modulation. The IOI values around the First Anomalous Accent (FAO) for the *rhythm* condition for those stimuli with sixteenth-note subdivision of the beat are given in Figures 4.7a and 4.7b. The IOI values around the FAO for the *tempo* condition for subdivided stimuli are given in Figures 4.7c and 4.7d. For all stimuli in Experiment 3, the initial beat rate was consistently 120 bpm, giving

Figure 4.6a

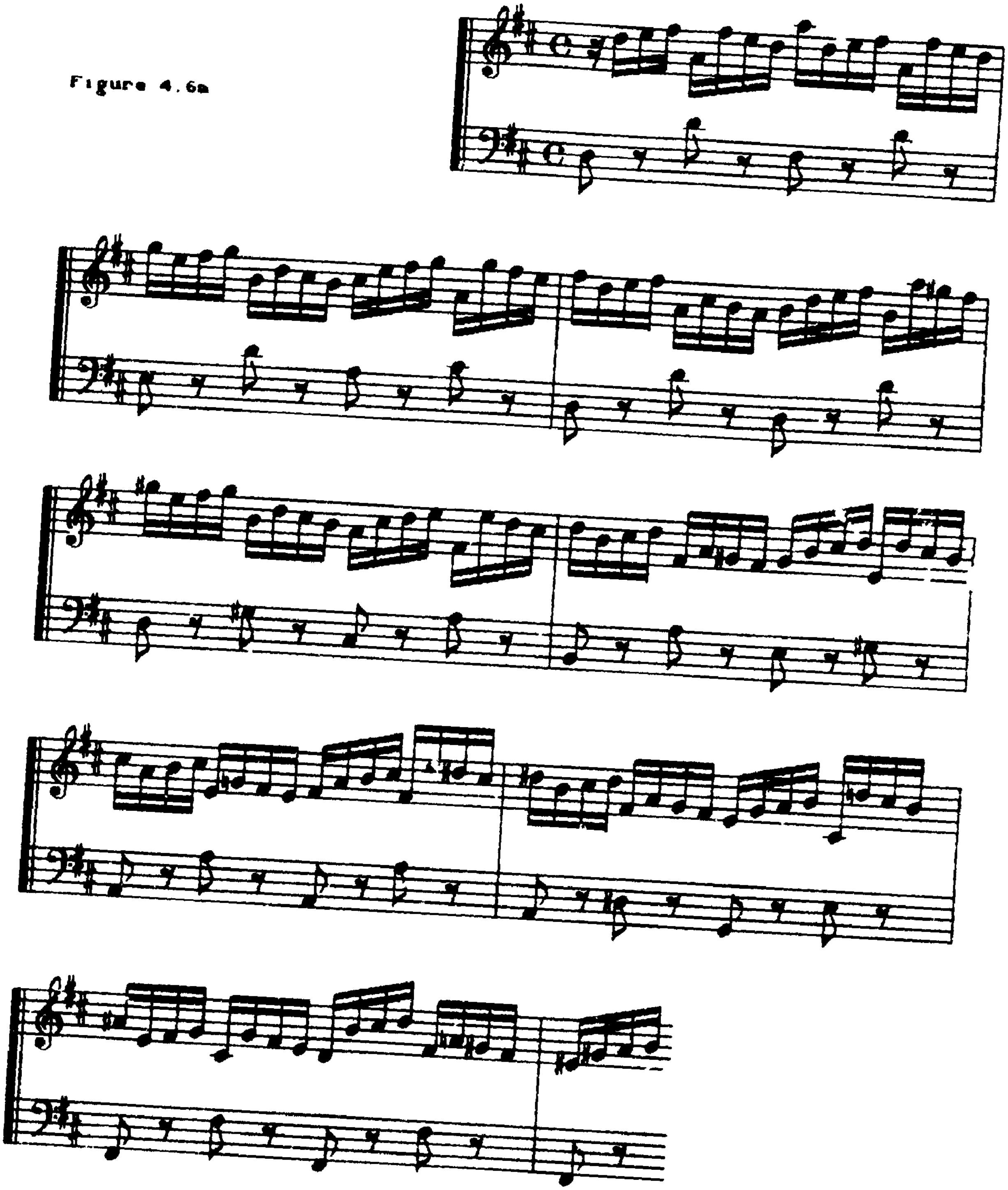
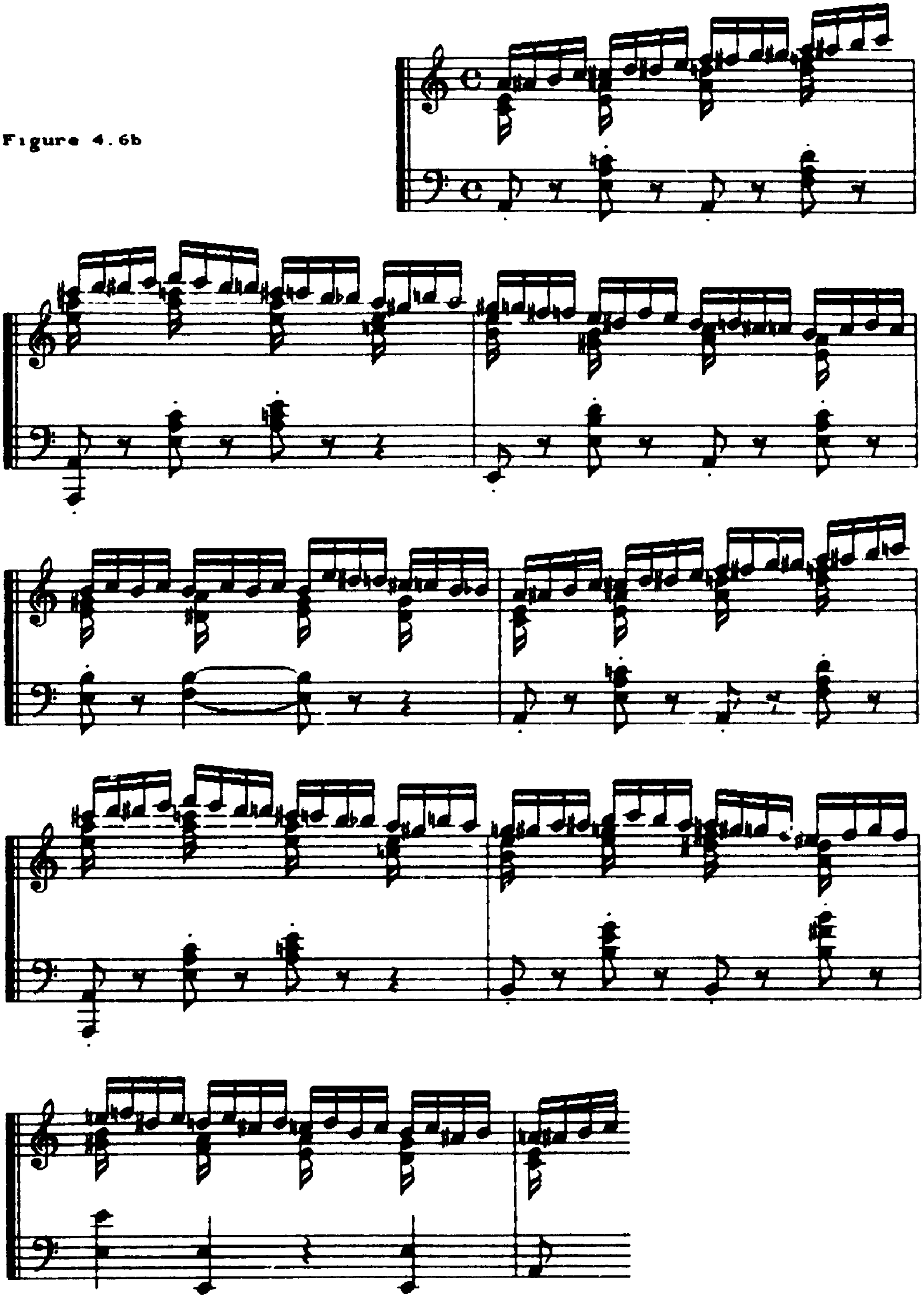



Figure 4.6b



data are understood. In classic reaction-time experiments, the subject is usually presented with a single stimulus event that he/she then has to categorize as belonging to one of two or more alternative sets (Pachella 1974; Lachman, Lachman and Butterfield 1979; Luce 1986). In this classic paradigm, the reaction time (in milliseconds) is taken as a measure of the amount of cognitive processing that is involved in placing the stimulus in one category or another. In other words, reaction time is assumed to be positively correlated with the cognitive complexity of the experimental task.

In contrast, in the present study, reaction time was used as a measure of the perceptual threshold for a particular type of modulation. Subjects were instructed to press the response button as soon as they detected a modulation, and NOT to attempt to identify the type of change at this stage. Thus, although subjects would eventually be expected to categorize each trial as either "rhythm change," "tempo change" or "constant," they did not have to make this decision until after the response button had been pressed. However, as Lachman et al (1979) and Pachella (1974) point out, one can never be sure that a simple detection task is entirely devoid of cognitive or decision-making content. The strategy of asking subjects to wait until the end of the trial to categorize the type of modulation will be assessed further in Chapter 6.

Since beats-to-respond constitutes at least intervallic data, Factorial Repeated-Measures Analysis of Variance (ANOVA) was chosen as the principal test by which these data would be analyzed. To test for reliability, Pearson's Product-Moment Correlation coefficient was calculated between original and replication versions of each condition

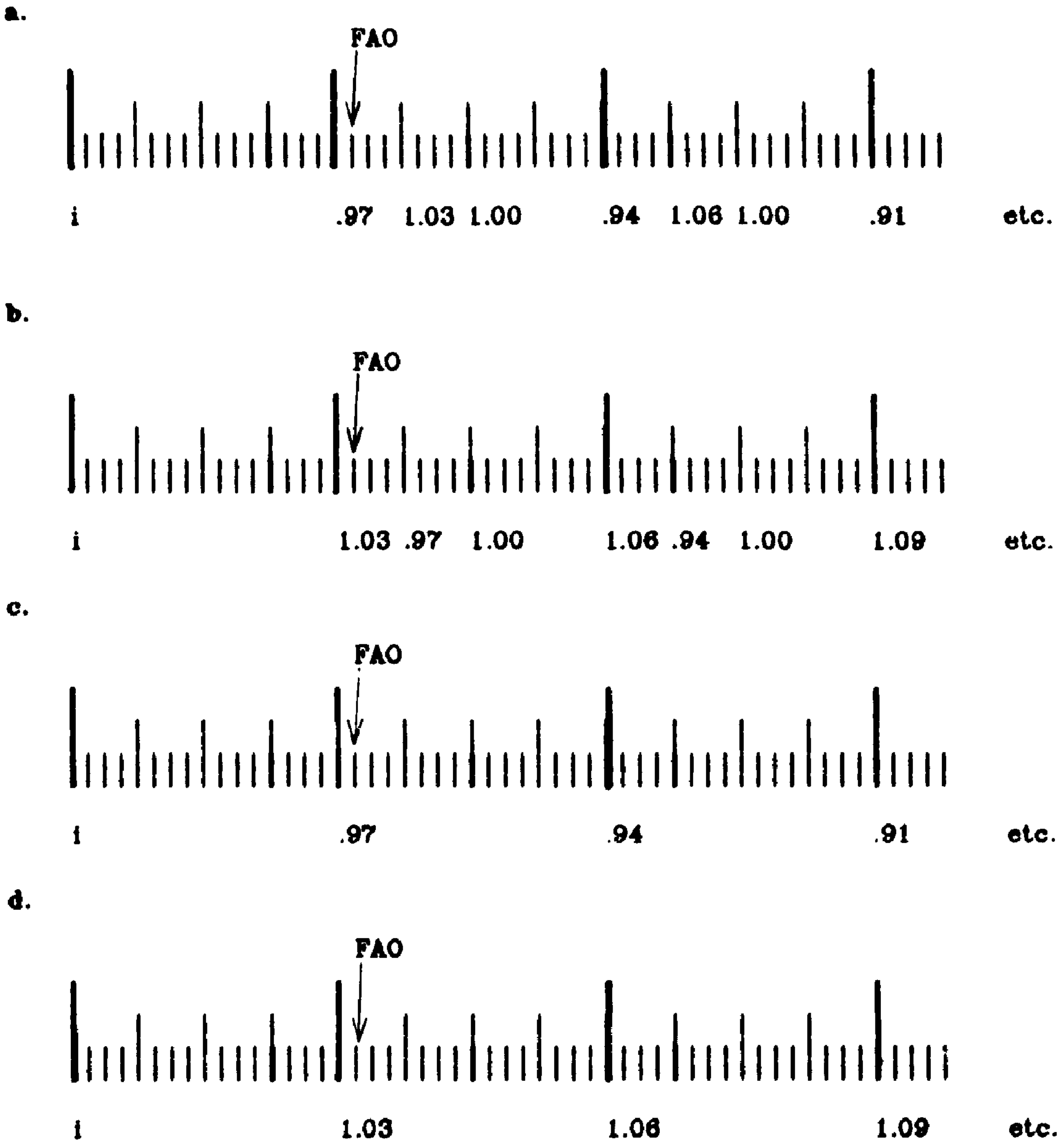


Figure 4.7. Generic stimuli for Experiment 3. Figure 4.7a shows the beginning of the rhythm-early stimulus. Figure 4.7b shows the rhythm-late stimlus. Figure 4.7c shows the tempo-early stimulus. Figure 4.7d shows the tempo-late stimulus. The numbers below the event onsets signifcy the same as they did in Figures 4.4 and 4.5. Note that each beat (quarter-note) time interval is divided equally into four sixteenths, regardless of its absolute time span.

i in Figure 4.7 a value of 500 ms. In addition, all anomalous onsets occurred on beat 2, as in Experiment 1.

DIRECTION. As in Experiments 1 and 2, DIRECTION had two levels, *early* and *late*, depending on when the first anomalous onset in each bar occurred relative to preceding beats. Also as in Experiment 1, the *early* condition involved IOIs getting shorter by 3% of the initial quarter-note IOI per bar, and the *late* condition involved IOIs getting longer by 3% of the initial quarter-note per bar.

PIECE. PIECE in Experiment 3 had four levels: *Bach*, *Chopin*, *Hindemith*, and *Author*. The *Bach*, *Chopin* and *Hindemith* pieces were chosen because they presented an isochronous sixteenth-note rhythm in the right-hand part, yet were in contrasting styles harmonically and melodically. The *Author* piece was composed to present an isochronous quarter-note rhythm: as such it was rhythmically identical to the stimuli of Experiment 1, but presented a tonal melodic and harmonic structure as well.

VERSION. As in Experiments 1 and 2, VERSION had two levels, *original* and *replication*. Again, the inclusion of two renditions of each condition allowed the calculation of a measure of reliability and helped determine whether subjects underwent fatigue or learning during trials.

The independent variables in Experiment 3 resulted in a 2 (TYPE) by 2 (DIRECTION) by 4 (PIECE) by 2 (VERSION) factorial design, giving 32 trials. In addition, one constant trial was included for each piece, giving 36 trials per subject. Again, constant trials were included merely as a foil, and data from them were disregarded in the statistical analysis.

Experiment 3 was designed to test Hypotheses 1 and 4, which may be re-expressed in terms of the above independent variables:

- Hypothesis 1. There will be a statistically significant interaction between TYPE and DIRECTION. *Rhythm* modulation will be indifferent to direction of modulation (*early* versus *late*): *tempo* modulation will be more difficult with *early* than *late*.
- Hypothesis 4. There will be a statistically significant effect owing to PIECE in the *rhythm* conditions. Owing to the lack of beat subdivision, *rhythm* modulation in the *Author* piece will be more difficult to detect than equivalent modulation in the three pieces with beat subdivision. It is unknown how PIECE will affect *tempo* modulation perception.

Dependent Variables

The purpose of all three experiments was to measure the relative difficulty of detecting rhythm and tempo modulation in the various contexts prescribed by the independent variables outlined above. In all three experiments, two dependent variables were used that reflected difficulty in detecting an IOI anomaly: (1) beats-to-respond data and (2) right/wrong answer data. Each dependent variable is considered in turn.

Beats-to-Respond Data

Definitions. The first measure of perceptual difficulty used was beats-to-respond, a special form of reaction-time data. A second glance at Figures 4.4 through 4.7 will confirm that the present study used ongoing, changing stimuli in which modulation was more and more pronounced with every bar. Given the nature of the stimuli, it is assumed that the more beats that were taken by listeners on average to detect a particular modulation, the more pronounced the IOI anomaly had to be to be detected. It is also assumed that the more pronounced an anomaly had to be before it was detected, the more difficult it was to hear that type of anomaly.

The average number of beats needed to detect a modulation was measured in the following way: For each trial an inaudible trigger pulse was generated that coincided exactly with the first anomalous onset (the beginning of the modulation). This trigger signal started a digital timer calibrated in milliseconds. When the subject thought that he/she had detected a modulation, he/she pushed a button that stopped the timer. The delay in milliseconds between the physical and perceived start of the modulation was recorded by the experimenter. At the end of a group of sessions, the reaction-time data in milliseconds were entered onto a computer and converted into beats-to-respond data. The process of data conversion is exemplified in Appendix E.

However, the use of reaction-time data gives rise to a number of issues that need to be addressed. Firstly, in order to be certain that reaction-time is indeed convertible to beats-to-respond data, one must consider what might be the component times of a typical datum value

(Pachella 1974; Welford 1980). For the present study, it was assumed that reaction time could be divided into two principal components: (1) detection time and (2) response time. Detection time and response time are illustrated in Figure 4.8 (see also Welford 1980).

Detection time is the time taken for a subject to notice an IOI anomaly. Since, as mentioned previously, the IOI anomalies get increasingly more pronounced by a fixed percentage of the initial IOI per bar, the physical extent of temporal change at any given time will depend on the number of beats that have elapsed since the start of the modulation. Thus the detection time is the time taken for an IOI anomaly to become pronounced enough to cross a perceptual threshold for a given subject; as such it is dependent on the rate at which stimulus events occur. Detection time is the value of interest in this investigation.

Response time is simply the time it takes a subject to press the response button. This component is determined by the subject's typical motoric speed as well as his/her state of arousal at the time of the test. As such, response time is independent of the rate of occurrence of stimulus events and merely adds error variance to the overall reaction-time measure. For this study it was assumed (a) that response time would vary randomly across trials and subjects and therefore would not affect results in any systematic way, and (b) that it would comprise only a very small percentage of the overall reaction time (Welford 1980).

Why not use simple reaction time in milliseconds rather than beats-to-respond? From the above discussion of the components of

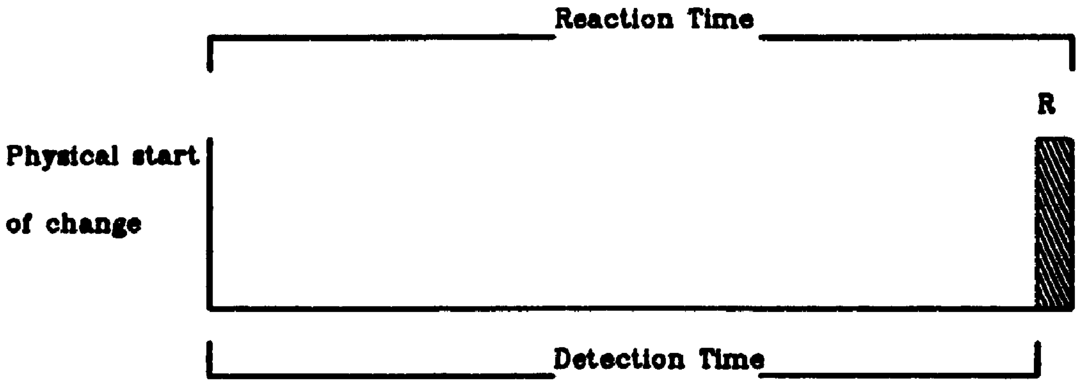


Figure 4.8. Assumed components of a reaction-time datum. The shaded area R represents the response time (the motoric delay), concluded by the closing of the switch that turns off the timer.

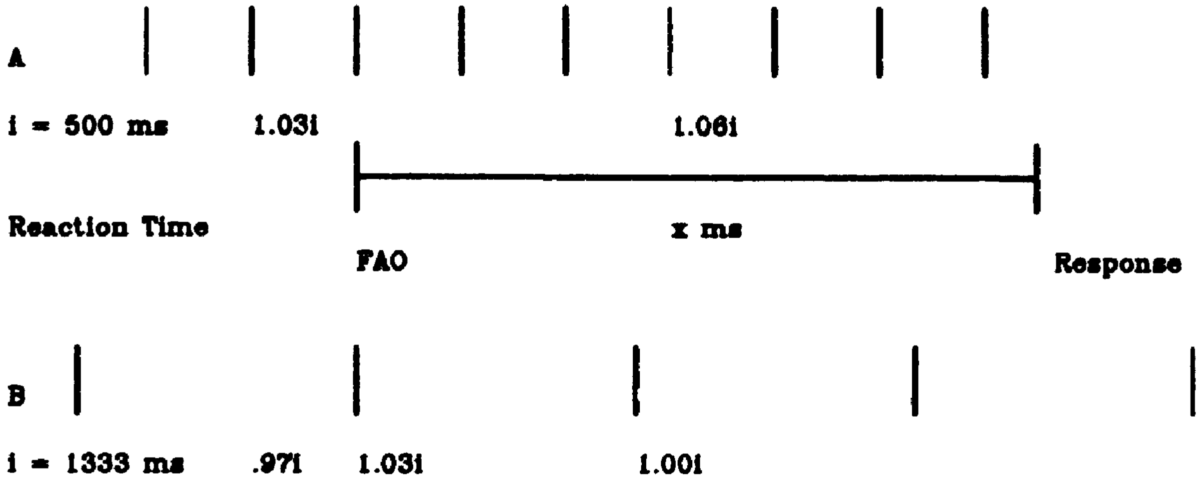


Figure 4.9. Single reaction time matched to two contrasting stimuli. Depending on the INITIAL BEAT RATE of the stimulus, one reaction-time datum can represent different numbers of beats elapsed.

reaction time, it can be seen that the major component, detection time, will depend on the number of stimulus beats that elapse in a given time. This dependence means that reaction times in milliseconds have to be interpreted in the light of the rate at which beats occur. To illustrate, Figure 4.9 shows the same reaction time matched to two different stimuli from Experiment 1. Stimulus A has an initial tempo of 120 bpm (500 ms) and then undergoes tempo modulation with the first anomalous onset coming late (getting slower). Stimulus B has an initial tempo of 45 bpm (1333.3 ms) and then undergoes a rhythm modulation with the first anomalous onset coming early. For stimulus A the subject has reacted to a modulation after 6.5 beats, but for stimulus B the reaction came after only 2.52 beats. Clearly, the total amount of percentage change of the initial IOI is different in the two cases, although the reaction time in milliseconds is identical. Therefore, for the purposes of measuring perceptual difficulty, it makes sense to convert millisecond values into the number of elapsed beats at the time of the response.[1]

From the above discussion, it is evident that the use of reaction-time data in the present study differed from the usual way such

---

[1] For the sake of rigour, one must also consider the somewhat unlikely null hypothesis that the number of beats elapsed has nothing to do with subjects' response, but instead that subjects are simply pressing the response button a fixed number of seconds after the beginning of the stimulus. If this were the case, seemingly statistically significant results would be obtained on beats-to-respond data purely from the data-conversion process. To check for this possibility, ANOVA tests were also carried out on raw reaction-time data in milliseconds. As will be seen from the brief discussion of the results of this check in Chapter 5, the foregoing null hypothesis (hereafter called the "data-conversion null hypothesis") could be rejected for all three experiments.

data are understood. In classic reaction-time experiments, the subject is usually presented with a single stimulus event that he/she then has to categorize as belonging to one of two or more alternative sets (Pachella 1974; Lachman, Lachman and Butterfield 1979; Luce 1986). In this classic paradigm, the reaction time (in milliseconds) is taken as a measure of the amount of cognitive processing that is involved in placing the stimulus in one category or another. In other words, reaction time is assumed to be positively correlated with the cognitive complexity of the experimental task.

In contrast, in the present study, reaction time was used as a measure of the perceptual threshold for a particular type of modulation. Subjects were instructed to press the response button as soon as they detected a modulation, and NOT to attempt to identify the type of change at this stage. Thus, although subjects would eventually be expected to categorize each trial as either "rhythm change," "tempo change" or "constant," they did not have to make this decision until after the response button had been pressed. However, as Lachman et al (1979) and Pachella (1974) point out, one can never be sure that a simple detection task is entirely devoid of cognitive or decision-making content. The strategy of asking subjects to wait until the end of the trial to categorize the type of modulation will be assessed further in Chapter 6.

Since beats-to-respond constitutes at least intervallic data, Factorial Repeated-Measures Analysis of Variance (ANOVA) was chosen as the principal test by which these data would be analyzed. To test for reliability, Pearson's Product-Moment Correlation coefficient was calculated between original and replication versions of each condition for each subject. The t-test was used to check whether there was any significant difference between mean response values for the two randomized orders in which the trials were presented, and for the two randomized positions of the First Anomalous Attack (bar two or bar three). In addition, a further subsidiary test was also required owing to one of the ubiquitous characteristics of reaction-time data, namely missing values. Consideration of this problem follows as the next subsection.

Missing Values. Missing values could occur for two reasons in the data gathered in this study: firstly if the subject failed to press the response button before the trial came to an end, that is, a "missed response;" secondly if the subject pressed the response button before a modulation had started, that is, a "premature response." In the case of a missed response the subject has not detected a modulation that actually occurred; in the case of a premature response, the subject has imagined a modulation before one actually occurred.

One solution would simply be to proceed with statistical analysis with the missing data points, resulting in unequal size cells in the ANOVA matrix. However, Keppel (1991, ch 13) notes two problems with unequal cells. Firstly, performing analysis with unequal cell sizes violates the assumptions on which ANOVA is based, if, in addition, the conditions of normality and homogeneity of variance are not met. Secondly and more seriously, if the missing data are not randomly distributed across cells, but are affected by particular experimental treatments, then the loss of data cannot simply be ignored. In the present study, the beats-to-respond data are intended to measure the relative difficulty of different types of temporal change. If, then, a subject fails to respond by the end of trial, this is surely an important indication that he/she found the modulation not only difficult to hear, but *imperceptible*. Therefore, omission of the missing data in this study would ignore a meaningful part of the subjects' responses.

Instead, all missing data points were converted to an arbitrary "maximally long" beats-to-respond value of 23. In all conditions, 23 was the total number of beats that elapsed from the first anomalous onset to the end of the bar that concluded the rhythm or tempo modulation. In those trials where the modulation began in bar 2, there was an extra bar after the modulation had been completed before the stimulus came to an end (see again Figures 4.4 and 4.5). Responses made during this bar (or indeed in the silence after the stimulus had finished), what might be termed "long reaction-time outliers," would be greater than 23 beats. All responses greater than 23 were also converted to 23, and treated as maximally long values along with missing values.

The conversion of missing values to a maximally long value of 23 is an imperfect solution. Such a procedure makes sense for missed responses since the fact that a subject failed to detect a modulation implies that he/she found modulation detection difficult for that particular condition. On the other hand, conversion is less satisfactory for premature responses. In this case, the simple fact that a subject imagined a modulation before one occurred, provides no information on how difficult the modulation would have been to detect once it had occurred. By the same token then, it would also be unsatisfactory to convert a premature response to a minimal beat-to-respond value of 0. In the end, premature responses were arbitrarily converted to 23. As it turned out, there was a remarkably low number of premature responses--only 11 responses across all three experiments out of a total of 2912 data points (or .4%). In addition these responses were randomly distributed across cells, suggesting that they had a negligible influence on results.

One overall problem with conversion of missing values and long outliers to a standard value of 23 still needs to be addressed. If the number of data points converted to a value of 23 within a specific condition is large, then conversion will produce a bimodal distribution, violating the assumption of normality. Keppel (1991, ch 5) shows how the F test at the heart of ANOVA is relatively uninfluenced by even substantial deviations from normality. However, to provide a check that missing values and long outliers could reasonably be converted to a maximally long value of 23, a further statistical test was performed, namely Survival Analysis.

Survival Analysis is a statistical test used in such fields as clinical medicine, actuarial science and quality-control engineering to compute in probabilistic terms the future lifetime of a patient policy holder or machine component. To illustrate Survival Analysis, consider a test intended to measure the durability of a sample of machine fan-belts. The fan-belts are tested without rest under extreme conditions and their time till breakdown (the dependent variable) is measured over a period of, say, five days. The breakdown times are recorded and from these data a probability density curve is drawn up showing the probability of a randomly-selected fan-belt lasting at least until (not breaking down before) any given time from the start of the test. A possible curve from this type of data might look something like the one illustrated in Figure 4.10a. However, such fan-belt testing would be complicated by the fact that, by the end of the test period, some fan-belts may not yet have broken down. If the test were to continue indefinitely they could be expected to break at some time in the future, but that event of interest, the moment of breakdown, is unknown at the conclusion of the test. Survival Analysis labels cases where the event of interest has not yet occurred as "censored cases." In effect, a second, dichotomous dependent variable is constructed from the continuous dependent variable of "time till breakdown." As illustrated in Figure 4.10b, the two values of the dichotomous variable (called the censoring variable) are simply "event-of-interest occurred" and "event of interest did not occur" (censored cases). The likelihood function of any randomly selected fan-belt surviving until a time t is then calculated as the product of the density function of failure at t and the probability of survival beyond the end of the experiment (the probability of a censored case) (Cox and Oakes 1984; Kalbfleisch and Prentice 1980).

Imagine now that the fan-belt test was designed to compare the durability of three different brands of fan-belt. In such a case the same time-till-breakdown data could be collected; however, three separate probability distributions (one for each fan-belt brand) would now be computed. Again, each probability distribution would take censored cases into account. Survival Analysis could then be used to

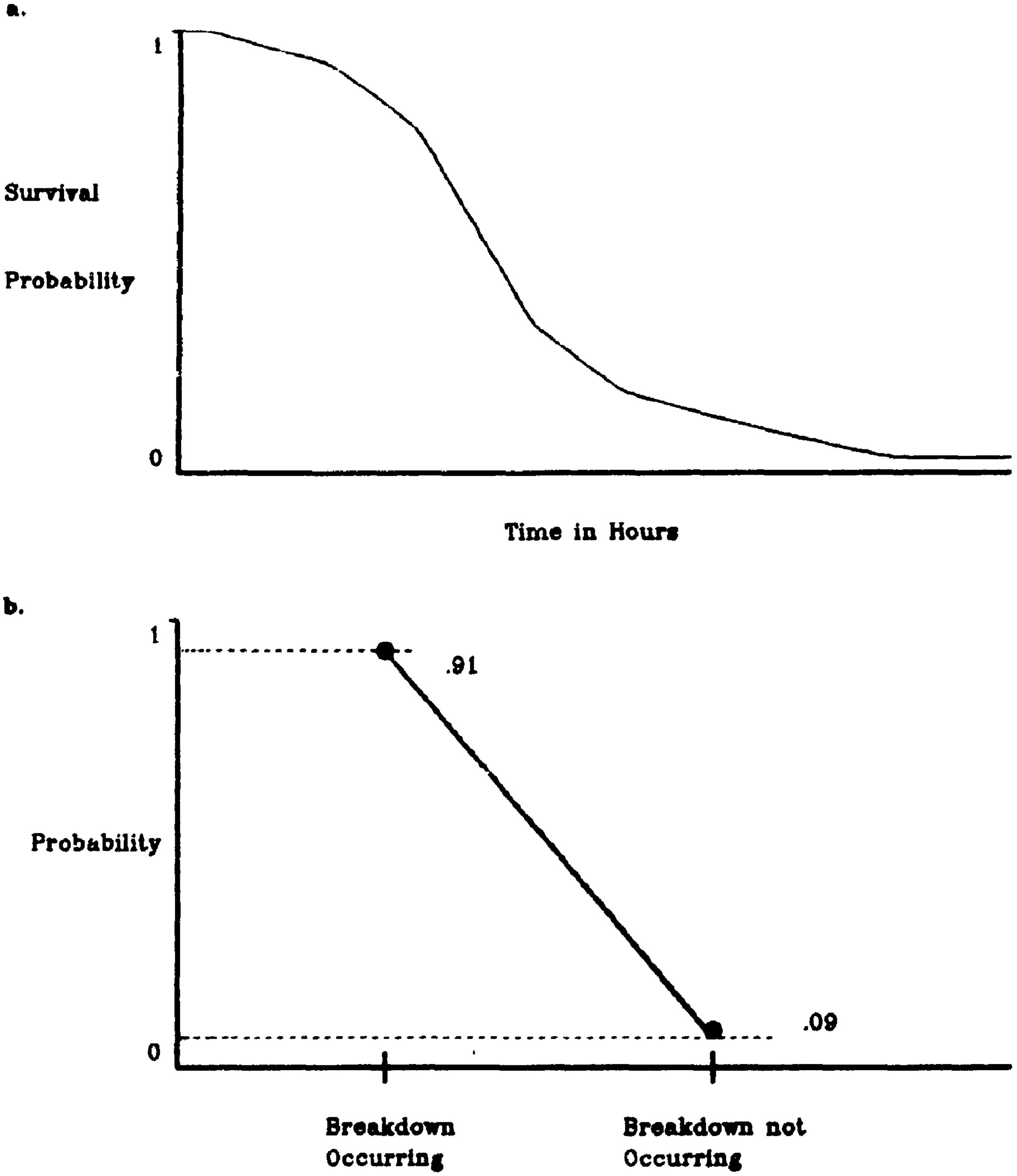


Figure 4.10. Probability curves used in Survival Analysis. Figure 4.10a shows the probability of survival of a fan-belt over time. Figure 4.10b shows the probability of a breakdown occurring during the testing period versus the probability of a breakdown not occurring during that period.

compute whether "brand of fan-belt" had significantly affected durability, including censored cases in that calculation. To do this, Survival Analysis would effectively compute a regression equation among the three survival-time distributions. The significance of the regression equation would give the significance of any association between fan-belt brand and durability.[2]

The fan-belt test situation may be readily translated to the present experiment. Instead of a test period of several days, the test period for each condition is 23 beats. Instead of different thicknesses of fan-belt, the independent variable can be any of the independent variables affecting subject detection of temporal change (TYPE, DIRECTION, INITIAL BEAT RATE and so forth). Instead of the breaking down of a fan-belt, the event of interest is the time point when the subject presses the response button. Correspondingly, censored cases will not be the absence of a fan-belt breakdown, but rather the absence of a time when the subject presses the response button (i.e. a premature or missing response represented by the value of 23).

Therefore, for the purposes of computing the Survival Analysis in the present study, the probability density function of a subject not responding until beat-number *b* was calculated for each level of each main effect using beats-to-respond data. Likewise, the probability of a no response occurring was computed for each level of each main effect using 23 as the censoring value. The degree of association between level of independent variable, and survival likelihood function was

---

[2] It should be stressed, however, that in calculating the regression equation, Survival Analysis uses the maximum likelihood method rather than the least-squares method (Kalbfleisch and Prentice 1980).

computed for each independent variable using the maximum likelihood statistic. Survival Analysis was therefore used in the present study to compute the significance of different independent variables on temporal-change detection whilst taking missing beats-to-respond data values into account. Therefore, if the same main effects were found to be significant in both the ANOVA test and the Survival Analysis, such a correspondence between the two tests would vindicate the conversion of long and missing values to an arbitrary value of 23 in the calculation of the ANOVA.

Identification Data.

The second dependent variable was identification data. In each trial, the subject was instructed to wait until the end of the stimulus and then identify which type of modulation had occurred. Thus, modulation-type identification was carried out after the subject had pressed the response button, and therefore constituted a separate response. Subjects could identify a given trial as an example of "rhythm change" (i.e., rhythm modulation), "tempo change" (i.e., tempo modulation) or "constant". The term "change" rather than "modulation" was used so far as the subjects were concerned; the experimenter wished to avoid any confusion that might have arisen from the tonal or harmonic connotations of the word "modulation." At the end of one experimental session, subjects' identifications were marked as either correct or incorrect and encoded on computer as 1s and 0s respectively.

For identification data, subjects had the opportunity to listen to the entire stimulus before they had to make a choice between "rhythm" "tempo" or "constant." Therefore two general strategies were possible on the part of subjects: (1) identify the type of modulation during the stimulus and hold this identification label in memory until the end, (2) wait until the end of the stimulus and then scan the entire stimulus in memory. The experiments described here had no way of telling which of the two strategies a given subject used. However, in any case, whichever means the subject used to identify modulation type, it was assumed that the proportion of correct versus incorrect answers for each cell would give a reliable and accurate measure of how difficult temporal change was to perceive given a particular combination of treatments. Thus, identification data were analyzed separately from, and in addition to, beats-to-respond data.

One fairly common procedure in much psychological experimentation is to treat right/wrong answer data (or proportion correct scores) as a continuous random variable, and therefore to analyze this type of dependent variable using ANOVA. Of course, binomial data (right versus wrong) will approach normality the larger the value of n, and therefore with a large sample there is no serious violation of the assumptions of ANOVA. However since the present study used four- and five-way factorial designs, individual cell sizes at the highest interaction level had an n of only 28 (the number of subjects used in each experiment). Using the following check:

$$p + (3\sqrt{(pq/n)}) \text{ must be } < 1$$

and

$$p - (3\sqrt{(pq/n)}) \text{ must be } > 0$$

(where p is the proportion correct responses and q is the proportion of incorrect responses), some of the cells did violate the assumption of normality whilst others did not (McClave and Dietrich 1991). It was therefore deemed wiser to treat the data as nominal and categorical. The test employed was a Log-Linear Analysis (Logit), with right versus wrong answers treated as the dependent variable. Log-Linear analysis is similar to the chi-square test in that it measures the likelihood that a given distribution across several discrete categories could occur by chance alone, thus assuming only nominal data. Unlike the two-way chi-square test, in Log-Linear Analysis two types of chi-square statistic were calculated: partial and marginal chi-square. For the partial chi-square, the expected frequency for each cell for each interaction was adjusted to allow for all other same- and lower-order interaction effects. For the marginal chi-square, the expected frequency for each cell in each two-way and higher interaction was computed ignoring the other interactions. Log-Linear Analysis was therefore able to handle multivariate designs such as the one used in this study (see Tabachnick and Fidell 1989, ch. 7).

Having discussed the general and experiment-specific features of the design of this study, the ensuing sections describe the logistics of preparing and running the three experiments, in short, the method. This description is divided up by experiment, although a number of features of the method are shared in common by all three experiments.

Part II - Method

Experiment 1

Preparation of Stimuli

All stimuli were generated on a microcomputer and recorded to high-fidelity chromium dioxide cassette tapes for presentation to subjects. The rhythmic and dynamic aspects of the stimuli were generated using the Performer 3.6 Sequencer software from Mark of the Unicorn on an Apple MacIntosh Plus Computer. This software apparently allowed definition of note durations down to 1 ms if desired, and 128 gradations of dynamic level. The sequencer drove a Yamaha DX7 II synthesizer set to an unpitched drum sound. The cassette machine used for recording was a Yamaha KX-800U Natural Sound Stereo Cassette Deck. In order to be able to check for the possibility that the order of trials affected subjects' performance, two random orders of trials (order "A" and order "B") were prepared on separate tapes.

Subjects

Twenty-eight undergraduate and graduate students from the Faculty of Music at the University of Western Ontario volunteered to participate as subjects in this study. Subjects were recruited through written advertisements, spoken announcements and word-of-mouth from various classes at the Faculty of Music. These volunteers were then randomly assigned either to trial order A or to trial order B.

Subjects were sampled from a population of university music students in order to reduce the possible between-subject variance. Since the dependent variables both measured the difficulty in hearing particular modulation types, such factors as subjects' musical training or ability might give rise to floor- or ceiling effects: subjects with no training/experience might find all modulations impossible to detect, whereas subjects with many years of training/experience might find all modulations too easy to detect. However, it was assumed that whilst training/experience would influence a subject's overall performance, it would not influence the trends among different treatment conditions.

Apparatus

Subjects were tested individually in a quiet, but not soundproofed room. The layout of the apparatus is shown in Figure 4.11. The experimental tape was played to subjects on a Yamaha KX-800U Natural Sound Stereo Cassette Deck through a pair of JVC HA-D330 Stereo Headphones. The output level of the tape deck was set to produce a comfortable and safe sound-pressure level. The stimulus was recorded on the left channel of the tape whilst the right channel was reserved for the trigger pulses that started the digital timer. The headphones were connected to the output of the tape deck in such a way that the left-hand (stimulus) channel sounded at equal volume in both ears of the headphones while the right-hand (trigger) channel was disconnected from them and therefore inaudible to the subject. Cross-talk between the two channels of the tape was minimized by recording the trigger signal at as low a level as possible. Instead, the right-hand channel was routed to the starting input of the digital timer, a Hewlett-Packard 5304A Timer/Counter linked to a Hewlett-Packard 5300 Measuring System. The display of the timer was hidden from the subject. The stopping input of

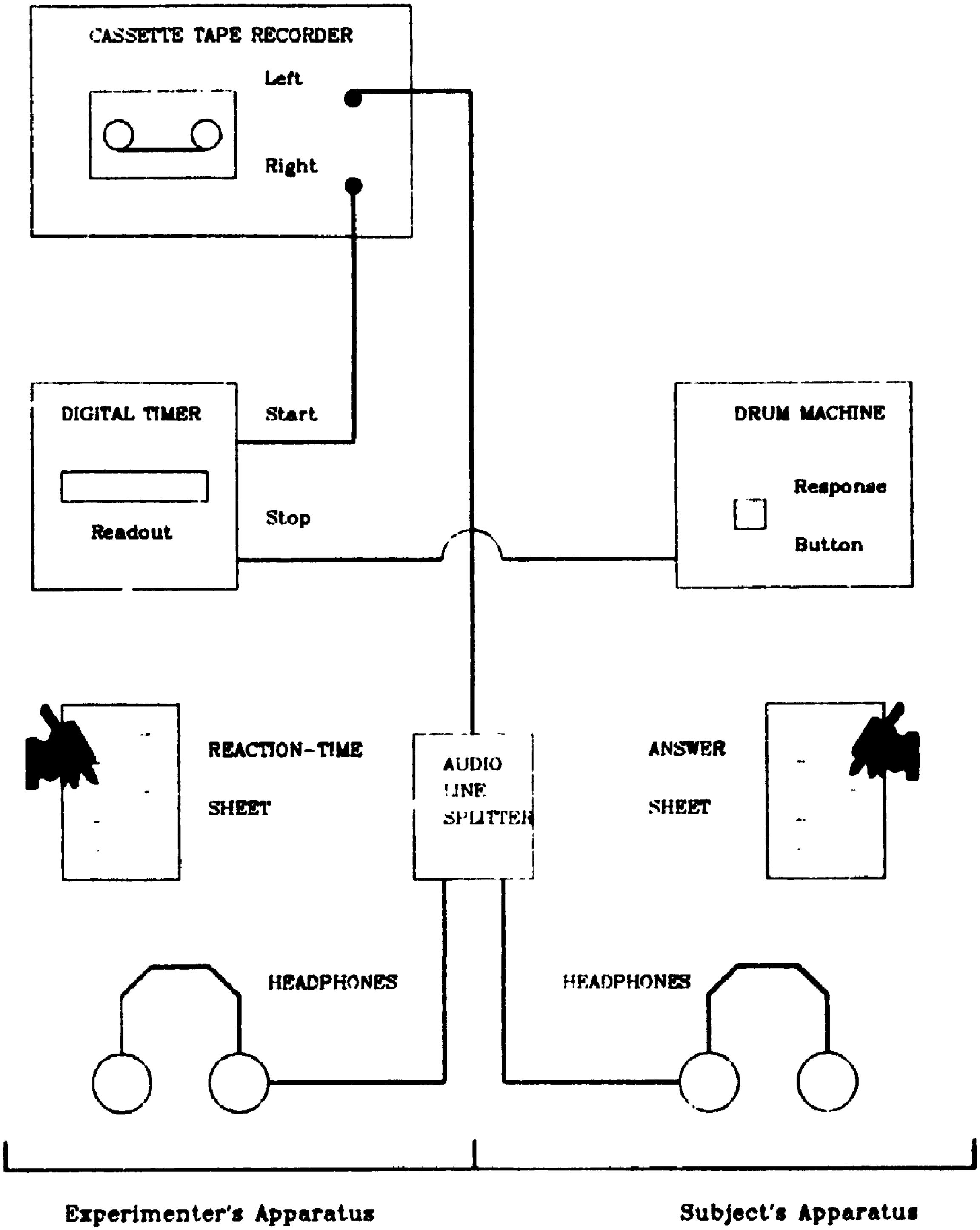


Figure 4.11. Set-up of experimental apparatus. The experimenter's apparatus was concealed from the subject throughout the experiment.

the timer was connected to a Roland TR-505 drum machine placed in front of the subject and set to "manual play." When the subject wished to respond, he/she pressed the "Snare Drum" button on the drum machine. The snare drum sound had an instantaneous attack and its signal caused the timer to stop as soon as the subject pressed the button. Lastly, on the table in front of the subject were also an answer sheet and a pen. To identify which type of modulation had occurred in a trial, the subject had to circle one of three letters on the answer sheet: "R" for "rhythm change," "T" for "tempo change" and "C" for "constant." The response sheets for all three experiments are given in Appendix D.

Procedure

On entering the room, the subject was given a sheet of written instructions (reproduced in Appendix D), and asked to read the details pertaining to the experimental task. While reading, the subject had the opportunity to ask the experimenter questions to clarify any points of procedure. Once the subject indicated comprehension of the experimental task, the stimulus tape (either order A or B) was started.

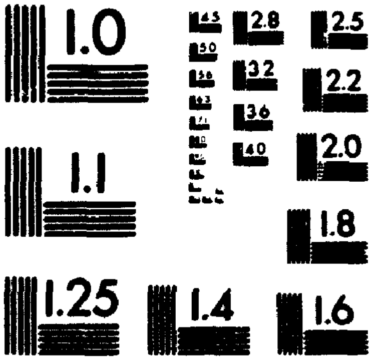
Both tapes began with the experimenter's voice thanking the subject and giving a brief recapitulation of the printed instructions. The first section of the tape consisted of 5 randomly-spaced, individual drum sounds that were separated by between 2 and 5 seconds. In response to these, the subject was asked to push the response button as soon as he/she heard a drum strike. In each case, the experimenter noted the response time on a concealed piece of paper and then reset the timer. The purpose of this was to test the subjects raw response time (the motoric component of reaction time) for subsequent checks in data analysis.

In the second section of the tape three practice trials were presented. The practice trials consisted of synthesized drum rhythms similar to those used as stimuli in the main part of the experiment (and as described in the "General Trial Format" for Experiment 1 above). These practice trials were intended to help the subject get used to interacting with the experimental apparatus (i.e., pressing the response button) as well as to exemplify the types of stimuli that would be presented. Immediately after each practice example, the subject was informed of the correct answer for modulation-type identification. Once all three practice trials had been played, the experimenter stopped the tape and asked the subject if everything was understood. If the subject understood the experimental task, the main part of the experiment proceeded. If the subject still indicated confusion, further explanation was given by the experimenter until the subject indicated comprehension.

The third section of the tape consisted of the 45 experimental trials in one of the two random orders. The procedure for each trial was as follows: (a) experimenter's voice announcing the trial number (b) the stimulus (as described above in Figures 4.4) during which the subject responded by pressing the response button as soon as he/she detected a modulation, (c) five seconds of silence during which the subject identified the modulation type by circling the appropriate letter on the answer sheet, and (d) five seconds of music on a mixture of vibraphone and electric organ sounds, whose pitches and rhythmic

3

PM-1 3½"x4" PHOTOGRAPHIC MICROCOPY TARGET
NBS 1010a ANSI/ISO #2 EQUIVALENT



PRECISION℠ RESOLUTION TARGETS

values were randomly generated on the Performer 3.61 sequencing program. The randomly generated music was intended as a "masking stimulus" to minimize memory of a previous trial influencing perception of the current trial.

The fourth and final section of the tape consisted of five more randomly spaced individual drum strikes. As in section one, these were intended to test the subject's raw response time. The purpose of introducing a second batch of response-time measurements was to see if the motoric component of the reaction time measurement slowed down or sped up over the course of the experimental session. The total length of the tapes for Experiment 1 was approximately 31 minutes.

At the end of the experimental session, the subject was thanked, paid $7 (Canadian), and debriefed. One session for Experiment 1, including time for reading instructions and debriefing was approximately 35 minutes.

Experiment 2

Preparation

The same equipment was used to prepare Experiment 2 as had been used in Experiment 1. The Yamaha DX7 II synthesizer was set to the same unpitched-drum voice and similar high-fidelity chromium dioxide tapes were used for recording purposes. As in Experiment 1, two different random order of the trials were prepared on two separate tapes.

Subjects

As in Experiment 1, 28 undergraduate and graduate music students from the University of Western Ontario volunteered to be subjects. The same recruitment procedures were used and these volunteers were then randomly assigned to one of the two trial orders. However, since learning had been observed over the course of Experiment 1, no subject took part in Experiment 2 who had taken part in Experiment 1.

Apparatus and Procedure

The room, the apparatus and the way in which each subject was asked to interact with the apparatus were identical to Experiment 1. Again, after reading the instructions, the subject first had to respond to five drum strikes to provide a measure of his/her motoric response time. Next, the subject heard three practice trials and then the trials proper. In Experiment 2, owing to the number of levels of each factor, there were only 36 experimental trials in one session. The format of one trial was the same as in Experiment 1, i.e.: (1) voice announcing trial number, (2) stimulus during which the subject responded by pressing the response button as soon as a modulation was detected, (3) 5 seconds silence during which the subject identified the modulation type by circling the appropriate letter on the answer sheet, and (4) 5 seconds of randomly generated music, similar to that used between trials in Experiment 1.

Finally, after the main trials, five more readings of the subject's basic motoric response time were taken. One session for

Experiment 2, including time for reading instructions and debriefing was approximately 30 minutes.

Experiment 3

Preparation

The same equipment was used to prepare Experiment 3 as had been used in Experiments 1 and 2. One difference was that the Yamaha DX7 II synthesizer was set to a grand piano sound rather than the unpitched-drum sound used in the previous experiments. Similar high-fidelity chromium dioxide tapes were used for recording purposes. As in Experiments 1 and 2, two different random orders of the trials were prepared on two separate tapes.

Subjects

As in Experiments 1 and 2, 28 undergraduate and graduate music students from the University of Western Ontario volunteered to be subjects. The same recruitment procedures were used as in the previous experiments with volunteers being randomly assigned to one of the two trial orders. However, since learning was observed in both of the previous experiments, no subjects who had taken part in either Experiment 1 or Experiment 2 were used in Experiment 3.

Apparatus and Procedure

The apparatus and the way in which each subject was asked to interact with it were identical to Experiments 1 and 2. After reading the instructions, the subject first had to provide a measure of his/her motoric response time. This time, however, in contrast to Experiments 1 and 2, the stimulus used to measure response time was a single synthesized piano tone at A = 440 Hz. A piano tone was used in Experiment 3 since this was more similar to the type of stimuli the subject was going to hear in the main experimental trials. As in the previous experiments, five randomly spaced tones were used. Next, the subject heard three practice trials and then the trials proper. In Experiment 3, owing to the number of levels of each factor, there were only 36 experimental trials in one session. The format of one trial was the same as in Experiment 1, i.e.: (1) voice announcing trial number, (2) stimulus during which the subject responded by pressing the response button as soon as he/she detected a modulation, (3) 5 seconds silence during which the subject identified the modulation type by circling the appropriate letter on the answer sheet, and (4) 5 seconds of randomly generated music, similar to that used between trials in Experiments 1 and 2.

Finally, after the main trials, five more readings of the subject's basic motoric response time were taken. One session for Experiment 3, including time for reading instructions and debriefing was approximately 30 minutes.

Conclusion

This chapter has outlined the hypotheses, the design, the preparation and the running of the three experiments that constitute the main part of this study. As should be evident, all three experiments share a number of features, including the general experimental tasks, the two types of dependent variable data collected, and certain independent variables (i.e. TYPE, DIRECTION and VERSION). The three experiments are distinguished by the use of other independent variables, namely INITIAL BEAT RATE, LOCATION and PIECE. The next chapter, Chapter 5, presents the results of these three related experiments, and, in so doing traces the commonalities and the differences in the way temporal change is detected in the contrasting contexts established by the various independent variables.

CHAPTER 5

EXPERIMENTAL RESULTS

This chapter describes the results of Experiments 1, 2 and 3. Some interpretation of the findings is offered; however, a more detailed discussion of how the results tie in with the model and previous literature is delayed until Chapter 6.

The relevant hypotheses and their null versions are reviewed for the sake of convenience at the beginning and at the end of the discussion of each experiment. For all three experiments, the alpha-level set to reject a particular null hypothesis was .01. This relatively conservative value was chosen in order to reduce the probability of a Type-I error (that is, accepting a difference between treatment conditions as "real" when it had in fact arisen by chance). Of course, at the same time, the probability of a Type-II error (that is, identifying a difference between treatment conditions as having arisen purely by chance, when it was in fact real) was thereby increased. Since the three experiments described here were designed to test a previously untested theoretical model, it seemed desirable to err on the side of caution. It will be remembered that, from Chapter 2, a review of the literature (particularly on tempo perception) revealed a bewildering tangle of contradictory results. The author did not intend to add to that confusion of data without being very sure that a true effect or interaction had in fact been observed.

Furthermore, as will be seen from the ensuing paragraphs, in addition to computing omnibus significances for main effects and interactions, a number of planned and *post hoc* comparisons between individual cell means or cell contingencies were carried out. The danger with multiple comparisons of this sort is an undesirable inflation of the cumulative probability of a Type-I error (Keppel 1991, ch. 8). Although one of the specific comparison tests used (Duncan) controlled for cumulative Type-I error probability within one set of comparisons, across the three experiments, several sets of comparisons were made, resulting in some inflation of the Type-I error probability. For this second reason, then, a stringent alpha-level of .01 seemed desirable.

Results for Experiment 1

Recall that Experiment 1 was intended primarily to test Hypotheses 1 and 2:

- Hypothesis 1. There will be a statistically significant interaction between TYPE and DIRECTION. *Rhythm* modulation will be indifferent to direction of modulation (*early* versus *late*); *tempo* modulation will be more difficult with *early* than with *late*.
- Null Hypothesis 1. There will be no statistically significant interaction between TYPE and DIRECTION. *Rhythm* and *tempo* modulation will exhibit the same trends *vis à vis early* and *late* anomalous onsets.

- Hypothesis 2. There will be a statistically significant interaction between TYPE and INITIAL BEAT RATE. *Rhythm* modulation will be most difficult to detect at extreme initial beat rates, whereas *tempo* modulation will be most difficult at slow initial beat rates.
- Null Hypothesis 2. There will be no statistically significant interaction between TYPE and INITIAL BEAT RATE. *Rhythm* and *tempo* modulation will exhibit essentially the same trends across the musical tempo range.

The results of the ANOVA performed on beats-to-respond data for Experiment 1 are summarized in Table III. The independent variables are abbreviated as follows: TYP = TYPE, DIR = DIRECTION, IBR = INITIAL BEAT RATE, VER = VERSION. The corresponding error term is given immediately below each source of variance. In each case the error term is the interaction between the source of variance in question and SUBJECTS, considered as a variable in its own right. Relevant statistics for the overall explained and residual sum of squares as well as the grand total sum of squares are given at the foot of Table III.[1]

---

[1] Also, in order to verify the validity of the results of Table III further, an ANOVA was performed on the raw reaction-time data in milliseconds from Experiment 1. This test was done to see if the "data-conversion null hypothesis" could be rejected (see Chapter 4, footnote 1). It was found that the same main effects and interactions were statistically significant for raw reaction-time as were for the beats-to-respond data. The one exception was the INITIAL BEAT RATE by VERSION interaction, which was not statistically significant for raw reaction-time data ($F = 1.45$, 4, 108 df, $p > .05$). The statistical significance of this interaction for beats-to-respond data should therefore be accepted with caution. In all other cases the data-conversion null hypothesis could be rejected and the results in Table III could be accepted.

TABLE III

EXPERIMENT 1. BEATS-TO-RESPOND DATA. ANALYSIS OF VARIANCE

Source of Variance	Sum of Squares	Df	Mean Square	F	p
TYP	772.45	1	772.45	19.64	.000
Error	1062.12	27	39.34		
DIR	1665.23	1	1665.23	37.24	.000
Error	1207.46	27	44.72		
IBR	4800.29	4	1200.07	53.09	.000
Error	2441.11	108	22.60		
VER	383.75	1	383.75	15.46	.000
Error	670.32	27	24.83		
TYP:DIR	3578.97	1	3578.97	138.57	.000
Error	697.34	27	25.83		
TYP:IBR	4754.33	4	1188.58	61.58	.000
Error	2084.51	108	19.30		
TYP:VER	109.26	1	109.26	9.24	.005
Error	319.36	27	11.83		
DIR:IBR	173.95	4	43.49	1.59	.182
Error	2952.97	108	27.34		
DIR:VER	8.13	1	8.13	.37	.547
Error	590.33	27	21.86		
IBR:VER	345.50	4	86.38	6.03	.000
Error	1548.18	108	14.33		
TYP:DIR:IBR	413.17	4	103.29	5.67	.000
Error	1968.07	108	18.22		
TYP:DIR:VER	1.55	1	1.55	.10	.754
Error	418.54	27	15.50		
TYP:IBR:VER	106.75	4	26.69	2.17	.077
Error	1326.32	108	12.28		
DIR:IBR:VER	187.87	4	46.97	3.97	.005
Error	1276.93	108	11.82		
TYP:DIR:IBR:VER	197.74	4	49.44	4.98	.001
Error					
Explained	17498.93	39	448.69	18.04	.000
Residual	26860.65	1080	24.87		
Total	44359.58	1119	39.64		

The results of the Survival Analysis are given in Table IV, which uses the same abbreviations as Table III. As can be seen from comparing Tables III and IV, the same main effects are found to be statistically significant using both ANOVA and Survival Analysis. Such an agreement between the two tests suggests that assigning the beats-to-respond value of 23 to missing responses was justified.

TABLE IV

EXPERIMENT 1, BEATS-TO-RESPOND DATA, SURVIVAL ANALYSIS

Source of Variance	p
TYP	< .0001
DIR	< .0001
IBR	< .0001
VER	< .0001

Table V shows the overall results of the Log-Linear analysis of the identification data for Experiment 1. Only those results that approach statistical significance (that is, $p < .1$) are shown on this table. One more abbreviation is introduced here: IDN stands for identification data.

TABLE V

EXPERIMENT 1. IDENTIFICATION DATA. LOG-LINEAR ANALYSIS

Dependent Variable	Independent Variable(s)	Df	Partial Chi-Squared	p	Marginal Chi-Squared	p
IDN:	TYP	1	40.789	.000	37.977	.000
IDN:	DIR	1	36.455	.000	33.821	.000
IDN:	IBR	4	47.173	.000	44.050	.000
IDN:	TYP:DIR	1	18.247	.000	17.297	.000
IDN:	TYP:IBR	4	45.080	.000	43.378	.000

The experimental hypotheses concerned themselves with interactions between factors. To see if the statistically significant interactions shown in Table III support the hypotheses, it is necessary to examine the cell means and cell contingency tables for each interaction separately. Only those interactions are discussed that are either of interest *ex hypothesi*, or that are statistically significant beyond the .01 level. Hereafter, for all tables showing comparison statistics between cell means and cell contingencies, those values joined by a square bracket are *not* statistically significantly different at the .01 level. The tests used to determine this statistical significance level were: Student's t-test between cell means in 2 by 2 tables, Duncan's multiple-range test between cell means in x by x tables, and Chi-square among cell contingencies for 2 by 2 and x by x tables. For all tables, the parenthesized numbers are the n for each cell (number of observations for each cell).

Hypothesis 1 predicts a statistically significant interaction between TYPE and DIRECTION. Table III reveals just such a statistically significant interaction for the beats-to-respond data (F - 138.57, 1, 27 df, p < .0005). Table VI gives the cell means of the beats-to-respond data for the TYPE by DIRECTION interaction; cell-comparison statistics were computed only between *early* and *late* conditions for this table. Figure 5.1 presents Table VI as a graph.

TABLE VI

EXPERIMENT 1, BEATS-TO-RESPOND DATA,
CELL MEANS OF 2-WAY INTERACTION,
TYPE BY DIRECTION

	DIRECTION	
TYPE	Early	Late
Rhythm	10.99 (280)	12.12 (280)
Tempo	16.22 (280)	10.21 (280)

Comparing cell means on Table VI shows that there was no statistically significant difference between *rhythm-early* and *rhythm-late* conditions (t = 2.26, 558 df, p > .01), whereas there was a statistically significant difference between *tempo-early* and *tempo-late* conditions (t = 12.34, 558 df, p < .01). Although *rhythm-late* was only marginally more difficult to detect than *rhythm-early*, *tempo-early* was statistically significantly more difficult than *tempo-late*. These results support the idea that the detection of rhythm modulation was largely independent of the direction of modulation, whereas, for tempo modulation, detection was much more difficult when the tempo accelerated. This result is consistent with Hypothesis 1.

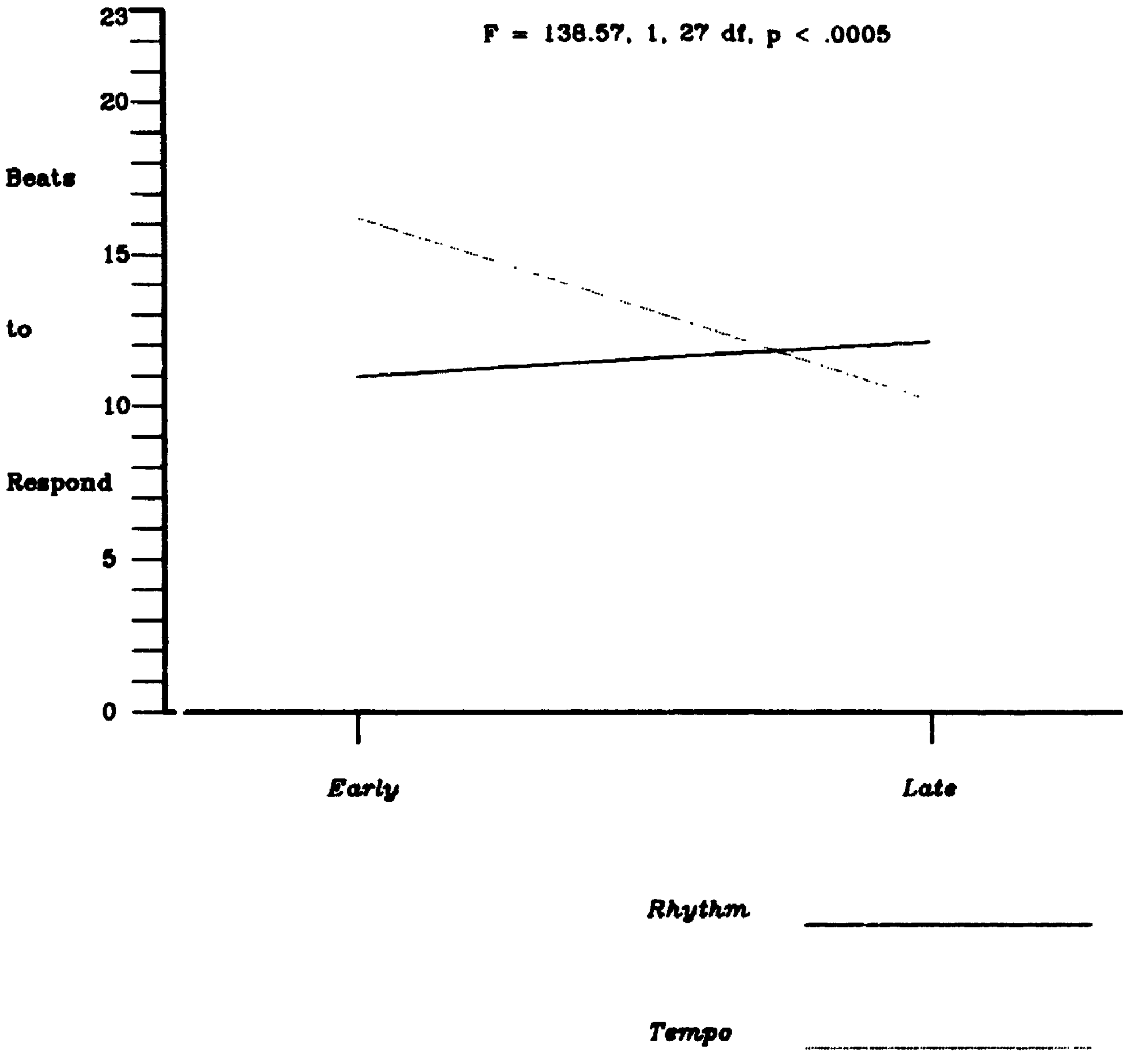


Figure 5.1. Graph of TYPE by DIRECTION interaction for beats-to-respond data in Experiment 1. The two levels of DIRECTION are shown on the x-axis.

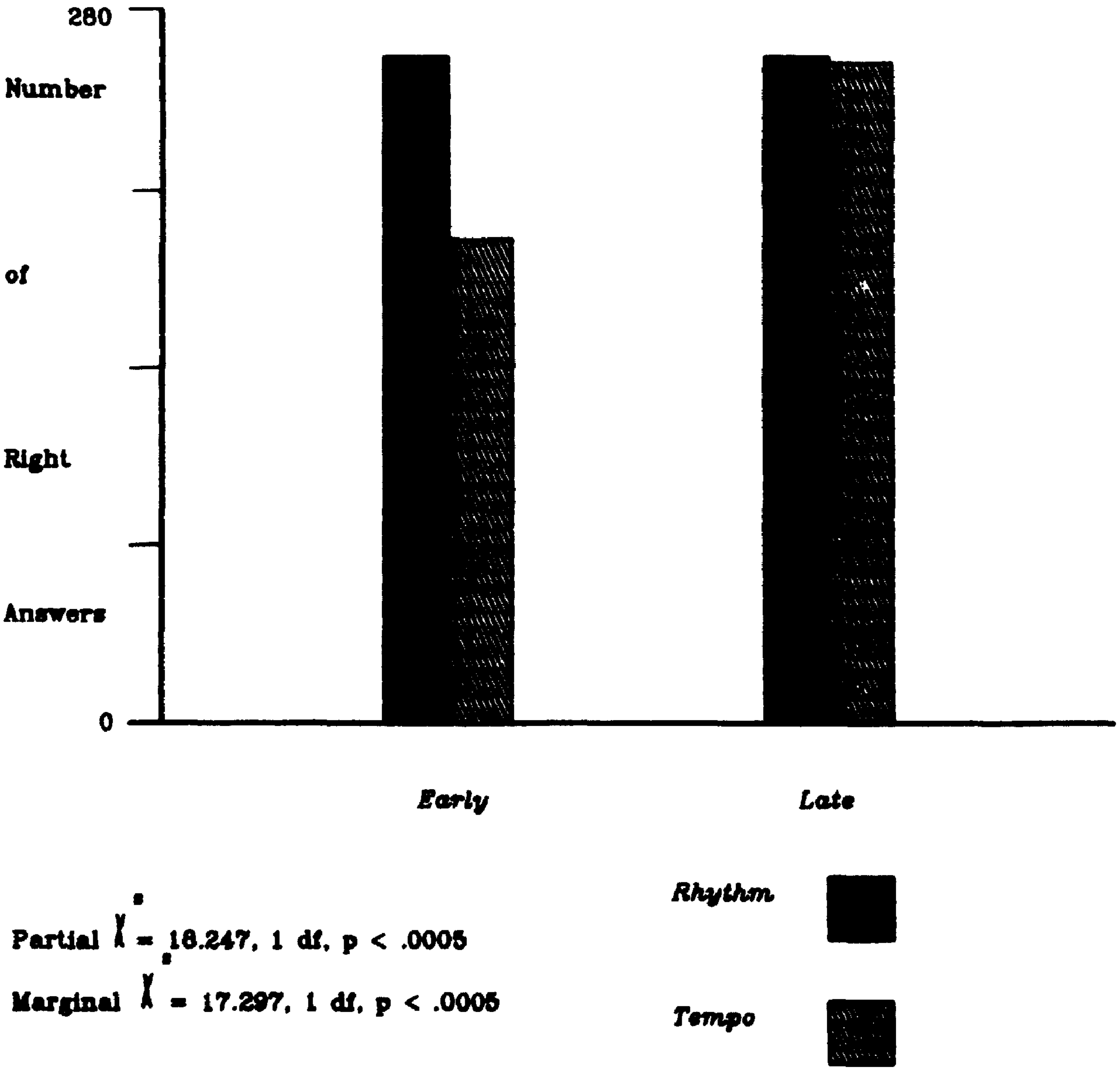
Hypothesis 1 is given further support by the identification data of Experiment 1. Table V shows a statistically significant interaction between TYPE and DIRECTION on the Log-Linear analysis (partial Chi-square = 18.25, 1 df, $p < .0005$; marginal Chi-square = 17.30, 1 df, $p < .0005$). Table VII, shows the tally of wrong and right answers, as well as the proportion correct, for the TYPE by DIRECTION interaction; cell-comparison statistics were computed only between *early* and *late* conditions for this table. Figure 5.2 presents Table VII as a histogram.

TABLE VII

EXPERIMENT 1, IDENTIFICATION DATA, CONTINGENCY TABLE FOR 2-WAY INTERACTION, TYPE BY DIRECTION

TYPE	DIRECTION	Wrong	Right	Proportion Correct	
Rhythm	Early	19	261	.932	(280)
	Late	19	261	.932	(280)
Tempo	Early	90	190	.679	(280)
	Late	21	259	.925	(280)

Table VII reveals the following: although exactly the same number of wrong versus right answers were recorded for the two directions of rhythm modulation (Chi-square = 0.00, 1 df, $p > .01$), for tempo modulation there were statistically significantly more wrong answers for the *tempo-early* condition, compared to the *tempo-late* condition (Chi-square = 51.96, 1 df, $p < .01$). Again, detection of rhythm modulation seemed to be independent of modulation direction.



Partial $\chi^2 = 18.247$, 1 df, $p < .0005$

Marginal $\chi^2 = 17.297$, 1 df, $p < .0005$

Figure 5.2. Histogram of TYPE by DIRECTION interaction for identification data in Experiment 1. The two levels of DIRECTION are shown on the x-axis. The total number of responses per cell is shown atop the y-axis.

whereas detection of tempo modulation was much more difficult when the tempo accelerated.

Hypothesis 2 predicts a statistically significant interaction between TYPE and INITIAL BEAT RATE. Table III reveals a highly significant interaction between TYPE and INITIAL BEAT RATE for the beats-to-respond data (F = 61.58, 1, 108 df, p < .0005). However, as evinced by Table VIII and Figure 5.3, the trend across INITIAL-BEAT-RATE levels for rhythm and tempo modulation predicted by Hypothesis 2 is only partially supported for this type of data. For Table VIII, cell-comparison statistics were computed between INITIAL-BEAT-RATE conditions but not between TYPE conditions.

TABLE VIII

EXPERIMENT 1, BEATS-TO-RESPOND DATA,
CELL MEANS OF 2-WAY INTERACTION, TYPE BY INITIAL BEAT RATE

	INITIAL BEAT RATE				
TYPE	45	75	120	190	300
Rhythm	8.36 (112)	8.24 (112)	9.52 (112)	12.90 (112)	18.76 (112)
Tempo	14.53 (112)	13.78 (112)	12.00 (112)	11.76 (112)	14.01 (112)

Hypothesis 2 proposed that detection of rhythm modulation would be most difficult at extreme initial beat rates, and easiest at moderate initial beat rates (taking 100 bpm as the mid-range value). In accordance with this hypothesis, the lowest mean (easiest to detect) is

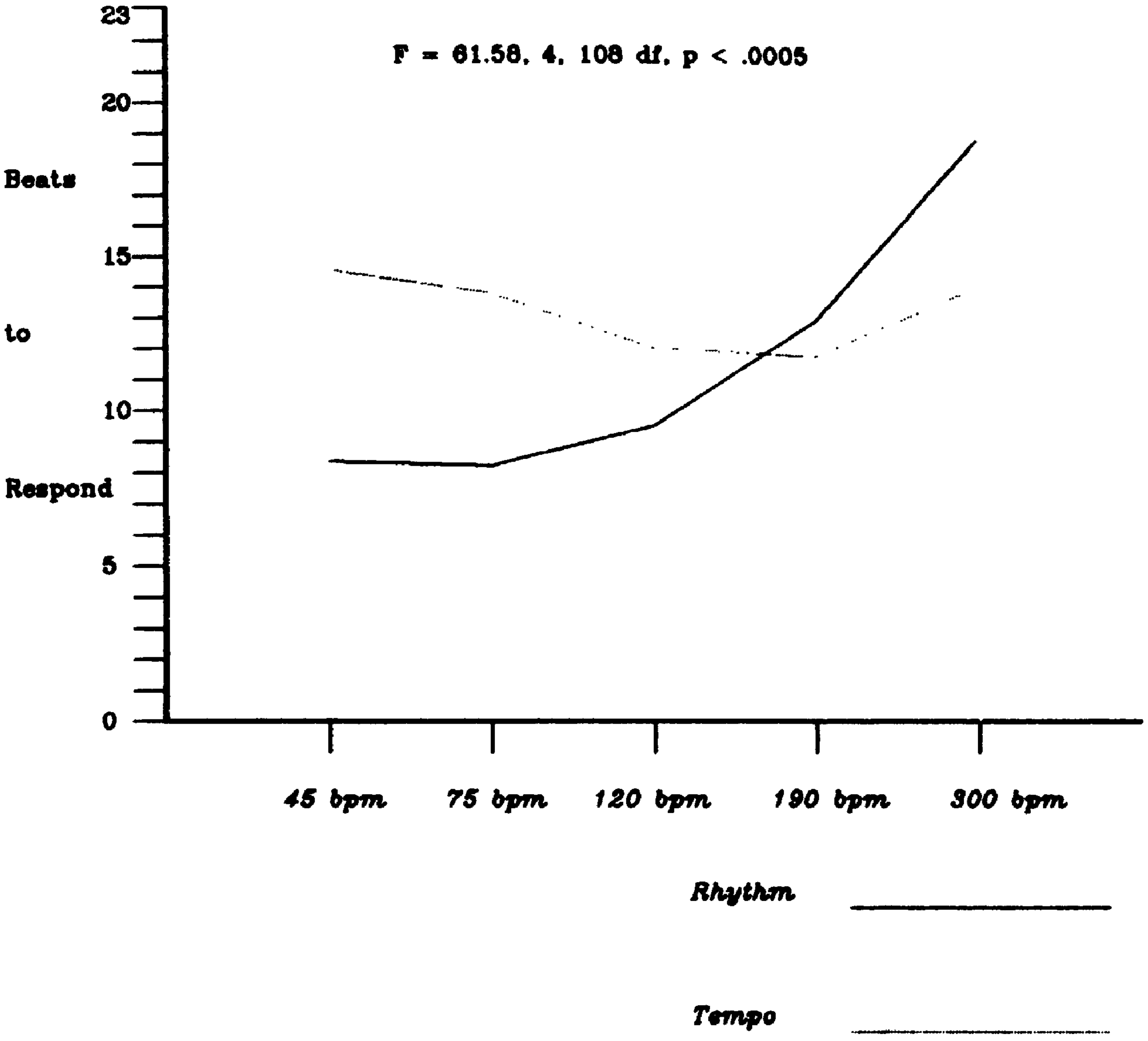


Figure 5.3. Graph of TYPE by INITIAL BEAT RATE interaction for beats-to-respond data. The five levels of INITIAL BEAT RATE are shown on the x-axis.

in fact the *rhythm-75* condition, and the highest mean (most difficult to detect) is the *rhythm-300* condition. In additional support of Hypothesis 2, *rhythm-120* is only marginally more difficult to detect than *rhythm-75* (Duncan p > .01). This is roughly as one might expect since 120 bpm is marginally nearer to the hypothetical mid-range tempo of 100 bpm on a logarithmic scale than 75 bpm is. However, contrary to Hypothesis 2, there is no statistically significant difference between the *rhythm-45* and *rhythm-75* cells (Duncan p > .01), although *rhythm-45* should be much more difficult as it is an extreme initial beat rate.

Hypothesis 2 also proposes that for tempo modulation, detection should get easier the faster the initial beat rate. This prediction is largely supported by the cell means in Table VIII, which show a gradual decrease in the beats needed to respond from *tempo-45* through *tempo-190*, with *tempo-45* being statistically significantly more difficult than *tempo-190* (Duncan p < .01). However, from the *tempo-190* to the *tempo-300* conditions there is a sudden increase in the beats-to-respond means. Whilst the cell-mean difference between *tempo-190* and *tempo-300* is not statistically significant at the .01 level (Duncan p > .01), this sudden increase is contrary to the trend across initial beat rates that exists up to the value of 190 bpm. Such a sudden break in the observed trend across initial beat rates for tempo modulation deserves further specific study.

Hypothesis 2 is supported more strongly by the identification data. Returning momentarily to Table V, one can see a statistically significant interaction between TYPE and INITIAL BEAT RATE using Log-Linear Analysis (Partial Chi-Square = 45.080, 4 df, p < .0005; Marginal

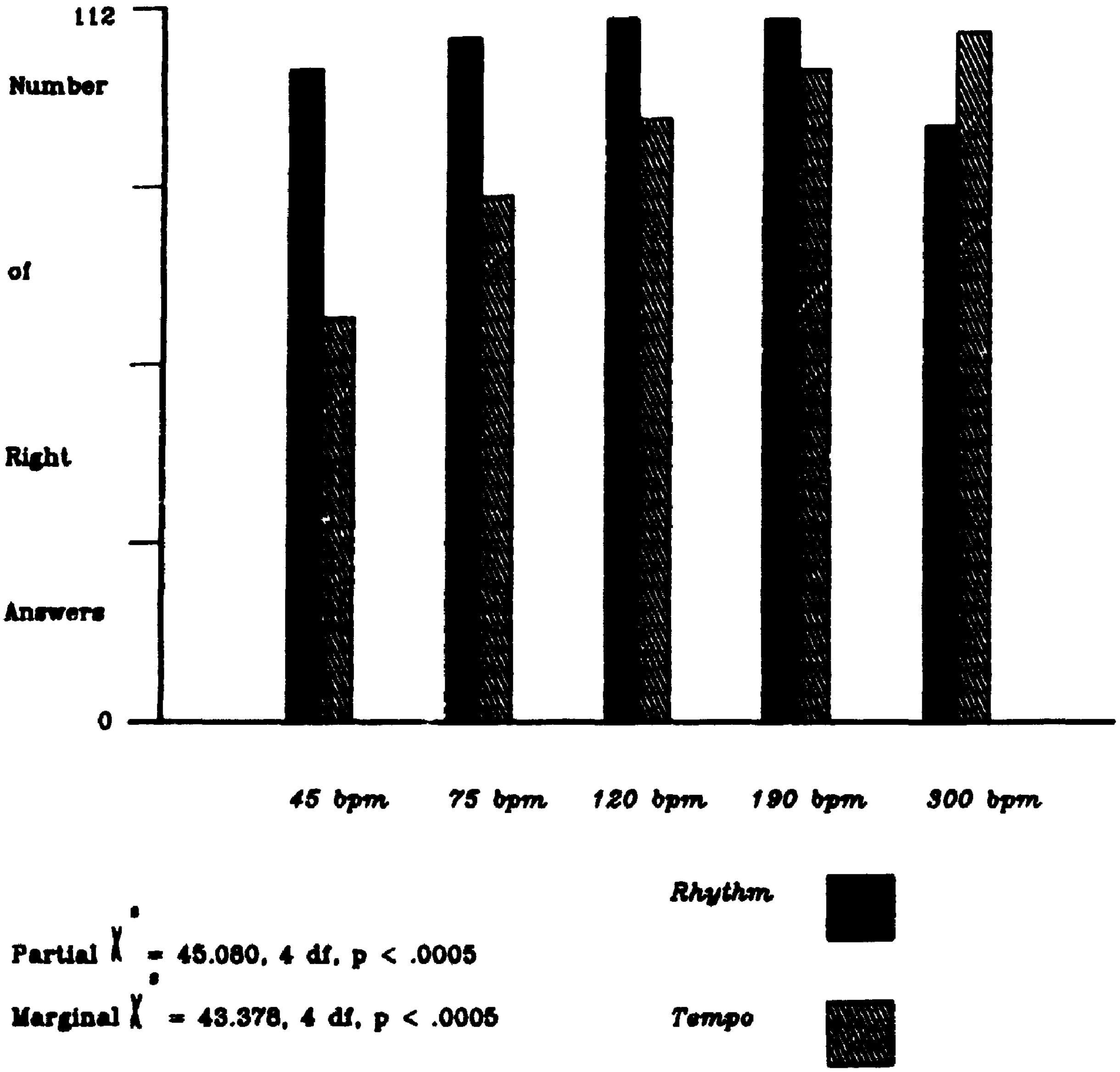
Chi-Square = 43.378, 4 df, p < .0005). The contingencies and the proportion-correct values for this interaction are given in Table IX and as a histogram in Figure 5.4. For Table IX, cell-comparison statistics were computed between INITIAL-BEAT-RATE levels but not between TYPE levels.

TABLE IX

EXPERIMENT 1, IDENTIFICATION DATA,
CONTINGENCY TABLE FOR 2-WAY INTERACTION,
TYPE BY INITIAL BEAT RATE

	TYPE					
	Rhythm			Tempo		
INITIAL BEAT RATE	Wrong	Right	Proportion Correct	Wrong	Right	Proportion Correct
45	10	102	.911 (112)	49	63	.563 (112)
75	5	107	.955 (112)	30	82	.732 (112)
120	2	110	.982 (112)	18	94	.839 (112)
190	2	110	.982 (112)	10	102	.911 (112)
300	19	93	.830 (112)	4	108	.964 (112)

In Table IX, rhythm modulation exhibits a slight drop in the numbers of wrong responses towards the centre of the initial-beat-rate range between 120 and 190 bpm. However, the only statistically significant difference between contingency cells is the difference between *rhythm-190* and *rhythm-300* (Chi-square = 8.11, 1 df, p < .01). Also, contrary to Hypothesis 2, the putative preferred initial-beat-rate



Partial χ^2 = 45.060, 4 df, p < .0005

Marginal χ^2 = 43.378, 4 df, p < .0005

Figure 5.4. Histogram of TYPE by INITIAL BEAT RATE interaction for identification data in Experiment 1. The four levels of INITIAL BEAT RATE are shown on the x-axis. The total number of responses per cell is shown atop the y-axis.

range shown in Table IX (*120-190 bpm*) is faster than the mid-range value of 100 bpm proposed in Chapter 3.

In support of Hypothesis 2, tempo modulation elicits fewer wrong answers, the faster the initial beat rate. This trend is maintained even for the *tempo-300* condition, which has the lowest number of wrong answers of all the tempo-modulation conditions. For tempo modulation, the only statistically non-significant differences are between adjacent INITIAL-BEAT-RATE levels (i.e., *tempo-45/tempo-75* Chi-square = 6.33, 1 df, p > .01; *tempo-75/tempo-120* Chi-square = 3.21, 1 df, p > .01; *tempo-120/tempo-190* Chi-square = 2.00, 1 df, p > .01; *tempo-190/tempo-300* Chi-square = 1.90, 1 df, p > .01). All other comparisons were statistically significant at the .01 level.

Both of the interactions, TYPE by DIRECTION and TYPE by INITIAL BEAT RATE, must be interpreted in light of the statistically significant three-way interaction TYPE by DIRECTION by INITIAL BEAT RATE for beats-to-respond data (F = 5.67, 4, 108 df, p < .0005) shown in Table III and below in terms of cell means in Table X and in Figure 5.5. For Table X, cell-comparison statistics were computed between DIRECTION levels and between INITIAL-BEAT-RATE levels, but not between TYPE levels.

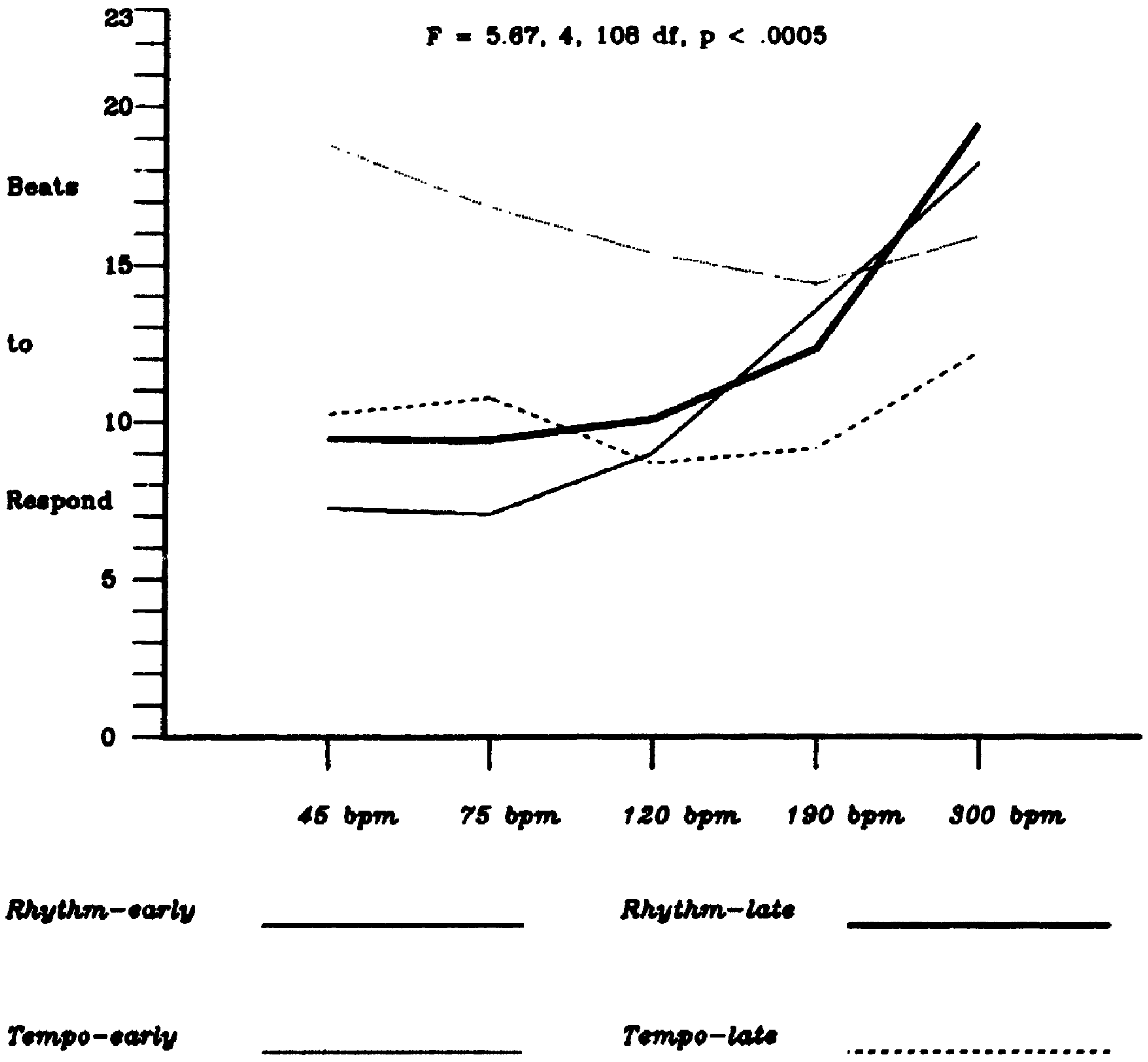


Figure 5.5. Graph of TYPE by DIRECTION by INITIAL BEAT RATE interaction for beats-to-respond data in Experiment 1. The Five levels of INITIAL BEAT RATE are shown on the x-axis.

TABLE X

EXPERIMENT 1, BEATS-TO-RESPOND DATA, CELL MEANS OF 3-WAY INTERACTION, TYPE BY DIRECTION BY INITIAL BEAT RATE

	TYPE			
	Rhythm		Tempo	
	DIRECTION		DIRECTION	
INITIAL BEAT RATE	Early	Late	Early	Late
45	7.26 (56)	9.47 (56)	18.78 (56)	10.27 (56)
75	7.05 (56)	9.43 (56)	16.80 (56)	10.77 (56)
120	8.99 (56)	10.06 (56)	15.33 (56)	8.68 (56)
190	13.50 (56)	12.30 (56)	14.38 (56)	9.14 (56)
300	18.15 (56)	19.37 (56)	15.84 (56)	12.19 (56)

In Table X, rhythm modulation follows essentially the same trend across the initial-beat-rate range for both *early* and *late* directions: rhythm-modulation detection gets statistically significantly more difficult with faster initial beat rates above 120 bpm (Duncan $p < .01$). Moreover, the only statistically significant difference between *rhythm-early* and *rhythm-late* is between *rhythm-early-75* and *rhythm-late-75* ($t = 3.52$, 110 df, $p < .01$).

For tempo modulation, *early* and *late* conditions exhibit different trends across the initial-beat-rate range and the difference between every *tempo-early/tempo-late* pair of means is statistically significant at the .01 level (according to the t-test). *Tempo-early* reveals a trend of gradually decreasing beats-to-respond values towards higher initial beat rates, with a slight increase in beats to respond between *190* and *300* bpm. However, only the differences *tempo-early-45/tempo-early-120* and *tempo-early-45/tempo-early-190* are statistically significant (Duncan $p < .01$). In contrast to *tempo-early*, *tempo-late* reveals a trend where there is a gradual decrease of beats-to-respond with 120 bpm eliciting the lowest response. However the only statistically significant differences across the initial-beat-range are *tempo-late-120/tempo-late-300* and *tempo-late-190/tempo-late-300*. This suggests that there may exist a perceptual threshold between 190 and 300 bpm when tempo is decelerating. This highly specific result is hard to interpret in the present context and may just be a Type-I error. Clearly more experimental work is needed to clarify this matter.

One further interaction in Experiment 1 deserves attention, namely INITIAL BEAT RATE by VERSION. Although this interaction was not predicted (and, as noted in footnote 1 above, the data-conversion null hypothesis could not be rejected for this interaction), it was found to be statistically significant for beats-to-respond data alone (from Table III, $F = 6.03$, 4, 108 df, $p < .0005$). Table XI shows the cell means for this interaction; for this Table, cell-comparison statistics were computed between VERSION levels only and not between INITIAL-BEAT-RATE levels.

TABLE XI

EXPERIMENT 1. BEATS-TO-RESPOND DATA.
CELL MEANS OF 2-WAY INTERACTION. INITIAL BEAT RATE BY VERSION

	INITIAL BEAT RATE				
VERSION	45	75	120	190	300
Original	11.16 (112)	11.39 (112)	11.64 (112)	13.73 (112)	16.94 (112)
Repli-cation	11.72 (112)	10.64 (112)	9.88 (112)	10.93 (112)	15.84 (112)

As can be seen from Table XI, the *replication* value is marginally lower (easier) than the *original* value for every initial beat rate except 45 bpm where it is slightly higher (more difficult). However, the only initial beat rate where the difference between *original* and *replication* is statistically significant is 190 bpm (t = 4.24, 222 df, p < .01). This highly specific result is hard to interpret in any theoretically meaningful way. Also difficult to interpret were the statistically significant three-way interaction DIRECTION by INITIAL BEAT RATE by VERSION (from Table III, F = 3.97, 4, 108 df, p < .01), and the four-way interaction TYPE by DIRECTION by INITIAL BEAT RATE by VERSION (F = 4.98, 4, 108 df, p < .005).

Finally with regard to Experiment 1, several other check statistics should be mentioned. For beats-to-respond data there was no effect owing to which of the two random orders the subjects heard (t = .4, 1118 df, p > .5), nor to starting point of the modulation within the trial (t = 1.36, 1118 df, p > .15). In addition, for identification data, there was no effect owing to order (Chi-square = .279, 1 df, $p > .5$) nor to start (Chi-square = .124, 1df, $p > .7$). These results are not surprising since the ordering of conditions and the placement of the FAO (bar 2 versus bar 3) were randomly determined.

The correlation coefficient for beats-to-respond data between *original* and *replication* versions of the same condition for each subject was $r = .63$ ($N = 560$, $p < .005$). This positive value is what one would expect, however, if $r = .63$, $r^2 = .39$, indicating that the *original* and *replication* versions only accounted for approximately 39% of each other's variance. The correlation coefficient for identification data between *original* and *replication* versions of the same condition for each subject was phi = .45 ($N = 560$, $p < .005$). Again, this positive value is generally what one would predict, however $r^2 = .20$, such that, on identification data, the two versions of each stimulus only account for about 20% of each other's variance. One explanation for these low r^2 values may be that while learning was apparently occurring across trials (VERSION as a main effect was statistically significant [$F = 15.46$, 1, 27 df, $p < .005$]), certain treatment conditions elicited more learning than others (the INITIAL BEAT RATE by VERSION interaction was statistically significant, for instance [$F = 6.03$, 4, 108 df, $p < .0005$]). Any interactions between VERSION and the other variables would mean that a trend that was observed between treatment conditions for the *original* version might not be observed for the *replication* version.

The correlation coefficient between the two dependent variables was in the direction expected (wrong answers correlating with longer beats-to-respond values) with $r = .43$ ($N = 1020$, $p < .005$). However, r^2

= .19, signifying that one dependent variable only accounted for approximately 19% of the variance in the other dependent variable. This meagre association may have arisen owing to the different time-points in a trial when the two types of data were collected, beats-to-respond during the trial and identification at the end. This discrepancy between the dependent variables will be discussed further in Chapter 6.

To conclude this description of the results for Experiment 1, one can summarize the status of the null hypotheses at this point:

- *Null Hypothesis 1. There will be no statistically significant interaction between TYPE and DIRECTION. Rhythm and tempo modulation will exhibit the same trends* vis à vis *early and late anomalous attacks.* According to the data from Experiment 1, Null Hypothesis 1 can be rejected. Both beats-to-respond data and identification data found a statistically significant interaction between TYPE and DIRECTION following the predicted trends.

- *Null Hypothesis 2. There will be no statistically significant interaction between TYPE and INITIAL BEAT RATE. Rhythm and tempo modulation will exhibit essentially the same trends across the musical tempo range.* According to the data from Experiment 1, Null Hypothesis 2 can be rejected. Both beats-to-respond and identification data found a statistically significant interaction between TYPE and INITIAL BEAT RATE. However, the precise trends predicted were only partially supported, and both types of data gave different trends across the initial-beat-rate range. These inconsistencies will be discussed further in Chapter 6.

Results for Experiment 2

Recall that Experiment 2 was intended primarily to test Hypotheses 1, 2 and 3:

- Hypothesis 1. There will be a statistically significant interaction between TYPE and DIRECTION. *Rhythm* modulation will be indifferent to direction of modulation (*early* versus *late*); *tempo* modulation will be more difficult with *early* than *late*.
- Null Hypothesis 1. There will be no statistically significant interaction between TYPE and DIRECTION. *Rhythm* and *tempo* modulation will exhibit the same trends *vis à vis early* and *late* anomalous attacks.
- Hypothesis 2. There will be a statistically significant interaction between TYPE and INITIAL BEAT RATE. *Rhythm* modulation will be most difficult to detect at extreme initial beat rates, whereas *tempo* modulation will be most difficult at slow initial beat rates.
- Null Hypothesis 2. There will be no statistically significant interaction between TYPE and INITIAL BEAT RATE. *Rhythm* and *tempo* modulation will exhibit essentially the same trends across the musical tempo range.
- Hypothesis 3. There will be a statistically significant interaction between TYPE and LOCATION. *Rhythm* modulation will be more difficult at *beat 1* than at *beat 2*; *tempo* modulation will be equally difficult to detect at *beats 1* and *2*.
- Null Hypothesis 3. There will be no statistically significant interaction between TYPE and LOCATION. *Rhythm* and *tempo* modulation

will exhibit the same trends across metric location of anomalous onsets.

The results of the Analysis of Variance (ANOVA) performed on beats-to-respond data for Experiment 2 are summarized in Table XII. The independent variables are abbreviated as follows: TYP = TYPE, DIR = DIRECTION, IBR = INITIAL BEAT RATE, LOC = LOCATION and VER = VERSION. The same layout is used for Table XII as was used for Table III.[2]

The results of the Survival Analysis are given in Table XIII, which uses the same abbreviations as Table XII. As can be seen from comparing Tables XII and XIII, the same main effects were found to be statistically significant using both ANOVA and Survival Analysis for all effects except LOCATION, which was found to be statistically non-significant by both tests. As in Experiment 1, such an agreement between the two tests suggests that assigning the beats-to-respond value of 23 to missing responses was justified.

---

[2] In order to test the data-conversion null hypothesis for Experiment 2, an ANOVA was also performed on the raw reaction-time data (in milliseconds). This analysis found that, generally, those main effects and interactions that were statistically significant for raw reaction-time data were the same as those that were statistically significant for beats-to-respond data. There were two exceptions: The two-way interaction DIRECTION by INITIAL BEAT RATE was statistically significant ($F = 19.42$, 1, 27 df, $p < .0005$) for raw reaction-time data when it was not statistically significant for beats-to-respond data. The three-way interaction TYPE by DIRECTION by INITIAL BEAT RATE was statistically significant ($F = 22.64$, 1, 27 df, $p < .0005$) for raw reaction-time data when it was not statistically significant for beats-to-respond data. Since these outstripped the significance of the corresponding results for beats-to-respond data, the data-conversion null hypothesis could be rejected in all cases for Experiment 2.

TABLE XII

EXPERIMENT 2, BEATS-TO-RESPOND DATA, ANALYSIS OF VARIANCE

Source of Variance	Sum of Squares	Df	Mean Square	F	p
TYP	754.59	1	754.59	17.03	.000
Error	1196.19	27	44.30		
DIR	3397.22	1	3397.22	56.20	.000
Error	1632.17	27	60.45		
IBR	227.61	1	227.61	6.85	.014
Error	896.52	27	33.20		
LOC	35.01	1	35.01	5.04	.033
Error	187.52	27	6.95		
VER	473.84	1	473.84	27.15	.000
Error	471.22	27	17.45		
TYP:DIR	3834.22	1	3834.22	61.63	.000
Error	1679.85	27	62.22		
TYP:IBR	820.35	1	820.35	24.73	.000
Error	895.54	27	33.17		
TYP:LOC	15.78	1	15.78	1.78	.194
Error	239.90	27	8.89		
TYP:VER	53.31	1	53.31	5.47	.027
Error	263.11	27	9.74		
DIR:IBR	66.60	1	66.60	2.73	.110
Error	657.61	27	24.36		
DIR:LOC	219.15	1	219.15	20.33	.000
Error	291.06	27	10.78		
DIR:VER	26.26	1	26.26	2.90	.100
Error	244.04	27	9.04		
IBR:LOC	11.86	1	11.86	1.06	.312
Error	302.04	27	11.19		
IBR:VER	64.58	1	64.58	3.40	.076
Error	513.45	27	19.02		
LOC:VER	16.19	1	16.19	1.40	.247
Error	311.82	27	11.55		
TYP:DIR:IBR	75.57	1	75.57	2.33	.139
Error	876.93	27	32.48		
TYP:DIR:LOC	49.28	1	49.28	2.89	.101
Error	461.01	27	17.07		
TYP:DIR:VER	219.30	1	219.30	10.99	.003
Error	538.78	27	19.95		
TYP:IBR:LOC	2.56	1	2.56	.45	.508
Error	153.77	27	5.70		
TYP:IBR:VER	32.94	1	32.94	4.33	.047
Error	205.36	27	7.61		
TYP:LOC:VER	82.91	1	82.91	4.11	.053
Error	544.54	27	20.17		

TABLE XII CONTD.

EXPERIMENT 2. BEATS-TO-RESPOND DATA. ANALYSIS OF VARIANCE

Source of Variance	Sum of Squares	Df	Mean Square	F	p
DIR:IBR:LOC	44.14	1	44.14	3.79	.062
Error	314.73	27	11.66		
DIR:IBR:VER	56.07	1	56.07	6.31	.018
Error	239.93	27	8.89		
DIR:LOC:VER	4.85	1	4.85	.57	.455
Error	227.93	27	8.44		
IBR:LOC:VER	30.32	1	30.32	1.71	.202
Error	478.69	27	17.73		
TYP:DIR:IBR:LOC	.88	1	.88	.05	.823
Error	467.17	27	17.30		
TYP:DIR:IBR:VER	7.83	1	7.83	.51	.481
Error	414.15	27	15.34		
TYP:DIR:LOC:VER	8.06	1	8.06	.50	.485
Error	434.51	27	16.09		
TYP:IBR:LOC:VER	21.59	1	21.59	1.62	.214
Error	359.73	27	13.32		
DIR:IBR:LOC:VER	66.61	1	66.61	4.35	.047
Error	413.66	27	15.32		
TYP:DIR:IBR:LOC:VER	31.23	1	31.23	2.71	.250
Error	310.93	27	11.52		
Explained	10750.71	31	346.80	14.74	.000
Residual	20325.70	864	23.53		
Total	31076.41	895	34.72		

TABLE XIII

EXPERIMENT 2. BEATS-TO-RESPOND DATA.
SURVIVAL ANALYSIS

Source of Variance	p
TYP	< .0001
DIR	< .0001
IBR	< .0001
LOC	> .2500
VER	< .0001

Table XIV shows the overall results of the Log-Linear analysis of the identification data. Only those results that approach statistical significance (that is, $p < .1$) are shown on this table. Again, IDN stands for identification data.

TABLE XIV

EXPERIMENT 2. IDENTIFICATION DATA. LOG-LINEAR ANALYSIS

Dependent Variable	Independent Variable(s)	Df	Partial Chi-Squared	p	Marginal Chi-Squared	p
IDN:	TYP	1	45.223	.000	41.772	.000
IDN:	DIR	1	31.130	.000	28.332	.000
IDN:	IBR	1	35.538	.000	32.495	.000
IDN:	LOC	1	5.367	.021	4.754	.029
IDN:	VER	1	7.301	.007	6.480	.011
IDN:	TYP:DIR	1	13.529	.000	11.542	.001
IDN:	TYP:IBR:VER	1	4.698	.030	4.951	.026

The hypotheses for Experiment 2 concerned themselves with interactions between factors. To see if the statistically significant interactions shown in Table XII support the hypotheses, it is necessary to examine the cell means and cell contingency tables for each interaction separately. Only those interactions are discussed that are either of interest *ex hypothesi*, or that are statistically significant beyond the .01 level. Again, for cell-comparison statistics, cell differences that are not statistically significant at the .01 level are indicated by a line linking the two cells in question.

Hypothesis 1 predicts a statistically significant interaction between TYPE and DIRECTION. Table XII reveals just such a statistically significant interaction for the beats-to-respond data ($F = 61.63$, 1, 27 df, $p < .0005$). Comparison of cell means for this interaction in Table XV (and Figure 5.6) shows that there was no statistically significant difference between *rhythm-early* and *rhythm-late* conditions ($t = .54$, 446 df, $p > .01$), whereas there was a statistically significant difference between *tempo-early* and *tempo-late* conditions ($t = 15.7$, 446 df, $p < .01$). For Table XV, cell-comparison statistics were only computed between *early* and *late* conditions.

TABLE XV

EXPERIMENT 2, BEATS-TO-RESPOND DATA, CELL MEANS OF 2-WAY INTERACTION, TYPE BY DIRECTION

	DIRECTION	
	Early	Late
TYPE		
Rhythm	11.38 (224)	11.62 (224)
Tempo	17.35 (224)	9.32 (224)

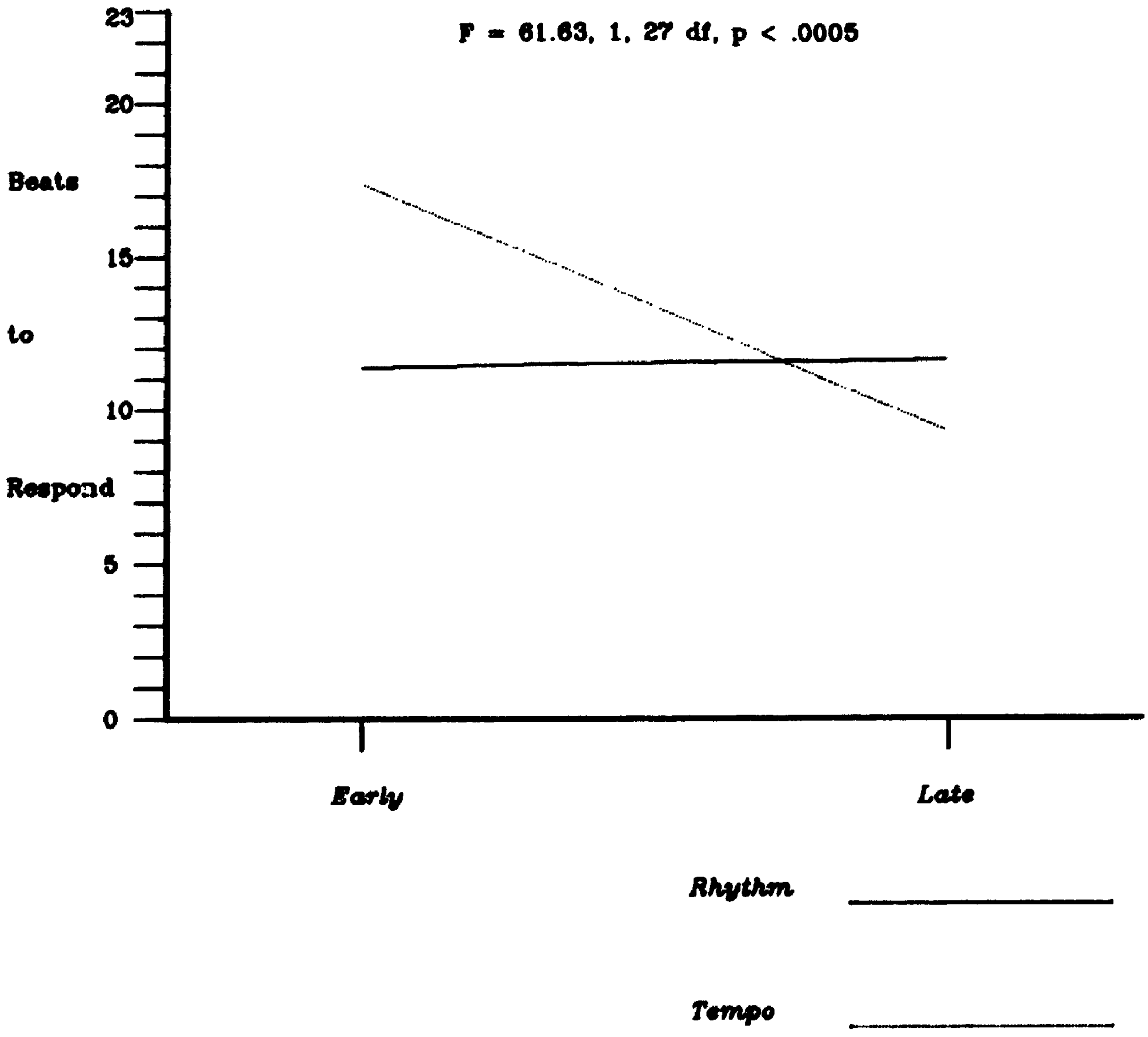


Figure 5.6. Graph of TYPE by DIRECTION interaction for beats-to-respond data in Experiment 2. The two levels of DIRECTION are shown on the x-axis

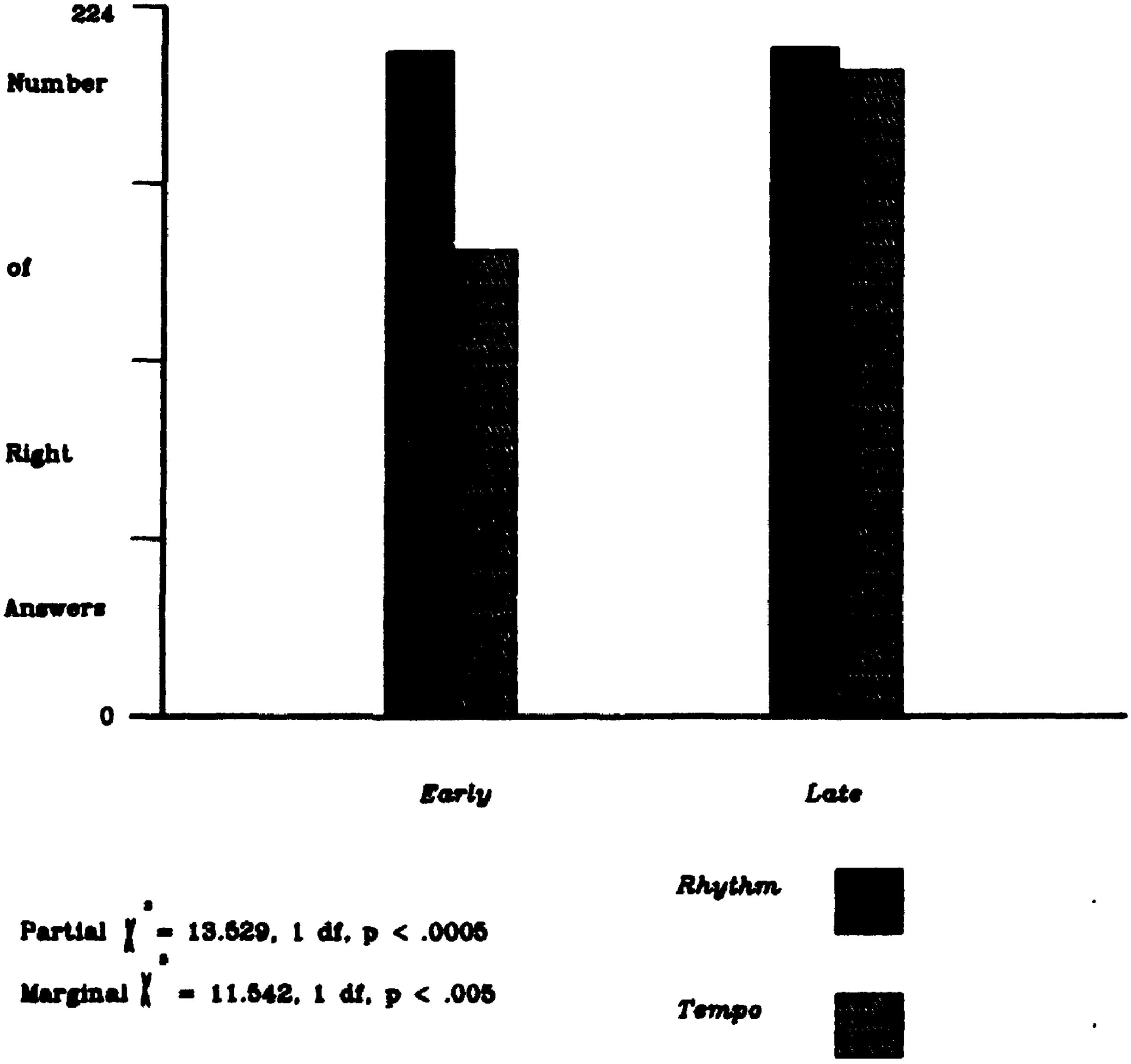
Although, in Table VI, *rhythm-late* was not statistically significantly more difficult to detect than *rhythm-early*, *tempo-early* was statistically significantly more difficult than *tempo-late*. Again, these results support the idea that the detection of rhythm modulation was largely independent of the direction of modulation, whereas, for tempo modulation, detection was much more difficult when the tempo accelerated.

Hypothesis 1 is given further support by the identification data of Experiment 2. Table XIV shows a statistically significant interaction between TYPE and DIRECTION on the Log-Linear analysis (partial Chi-square = 13.529, 1 df, $p < .0005$; marginal Chi-square = 11.542, 1 df, $p < .005$). Table XVI and Figure 5.7 give the tally of wrong and right answers and the proportion-correct values for the TYPE by DIRECTION interaction; cell-comparison statistics were only computed between DIRECTION conditions.

TABLE XVI

EXPERIMENT 2, IDENTIFICATION DATA,
CONTINGENCY TABLE FOR 2-WAY INTERACTION, TYPE BY DIRECTION

TYPE	DIRECTION	Wrong		Right	Proportion Correct
Rhythm	Early	15	I	209	.933 (n = 224)
	Late	14		210	.938 (n = 224)
Tempo	Early	78		146	.652 (n = 224)
	Late	21		203	.906 (n = 224)



Partial χ^2 = 13.529, 1 df, p < .0005

Marginal χ^2 = 11.542, 1 df, p < .005

Figure 5.7. Histogram of TYPE by DIRECTION interaction for identification data in Experiment 2. The two levels of DIRECTION are shown on the x–axis. The total number of responses per cell is shown atop the y–axis

Table XVI reveals the following: while almost the same number of wrong versus right answers was recorded for the two directions of rhythm modulation (Chi-square = 0.00, 1 df, p > .01), for tempo modulation there were statistically significantly more wrong answers for the *tempo-early* condition, compared to the *tempo-late condition* (Chi-square = 40.66, 1 df, p < .01). Again, detection of rhythm modulation seemed to be independent of modulation direction, whereas detection of tempo modulation was much more difficult when the tempo accelerated.

In the interests of completeness, the interaction TYPE by DIRECTION should be interpreted in the light of the three-way interaction TYPE by DIRECTION by VERSION, which was statistically significant on the beats-to-respond data alone (F = 10.99, 1, 27 df p < .005, according to Table XII). Table XVII gives the cell means for this 3-way interaction; for this table no cell-comparison statistics were computed.

TABLE XVII

EXPERIMENT 2, BEATS-TO-RESPOND DATA,
CELL MEANS OF 3-WAY INTERACTION, TYPE BY DIRECTION BY VERSION

	VERSION			
	Original		Replication	
	DIRECTION		DIRECTION	
TYPE	Early	Late	Early	Late
Rhythm	11.54 (112)	12.43 (112)	11.22 (112)	10.81 (112)
Tempo	18.99 (112)	9.62 (112)	15.71 (112)	9.01 (112)

Trends for rhythm and tempo modulation across the two different directions are the same for both the *original* and *replication* versions as they were in the two-way interaction TYPE by DIRECTION. However, for the *replication* version, the difference between *tempo-early* and *tempo-late* is less pronounced than the same difference in the *original* version. This 3-way interaction is of no theoretical interest and does not seem to alter the above interpretation of the TYPE by DIRECTION interaction.

Hypothesis 2 predicts a statistically significant interaction between TYPE and INITIAL BEAT RATE. Table XII reveals a statistically significant interaction between TYPE and INITIAL BEAT RATE for beats-to-respond data (F = 24.73, 1, 27 df, $p < .0005$). However, as can be seen from Table XVIII, which shows the cell means for the TYPE by INITIAL BEAT RATE interaction, the trend across INITIAL BEAT RATE levels predicted by Hypothesis 2 is not supported. For this table, cell-comparison statistics were computed only between INITIAL-BEAT-RATE conditions. Figure 5.8 reproduces Table XVIII as a graph.

TABLE XVIII

EXPERIMENT 2, BEATS-TO-RESPOND DATA,
CELL MEANS OF 2-WAY INTERACTION,
TYPE BY INITIAL BEAT RATE

	INITIAL BEAT RATE	
	60	150
TYPE		
Rhythm	10.04 (224)	12.96 (224)
Tempo	13.79 (224)	12.88 (224)

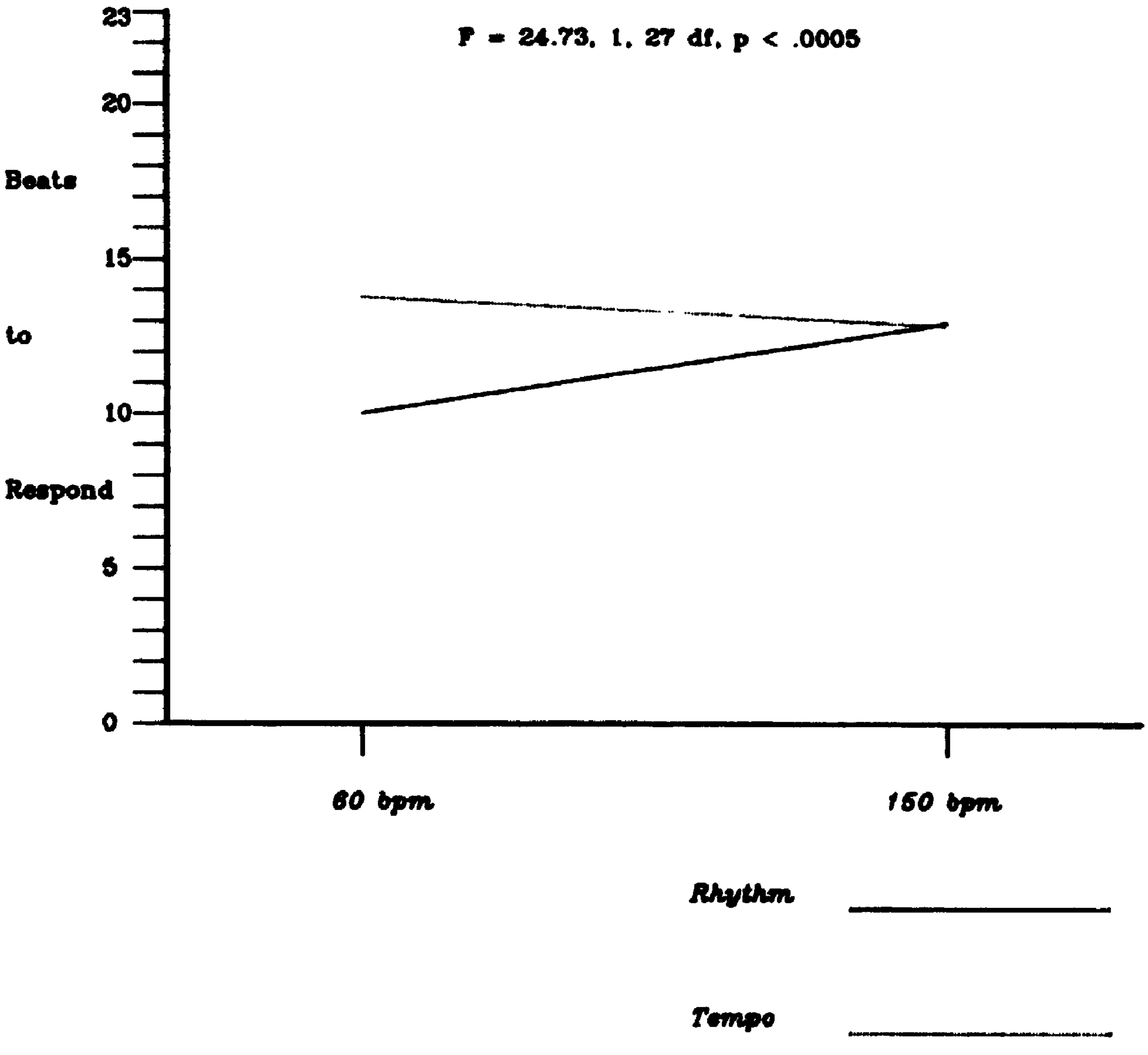


Figure 5.8. Graph of TYPE by INITIAL BEAT RATE interaction for beats-to-respond data in Experiment 2. The two levels of INITIAL BEAT RATE are shown on the x-axis.

Hypothesis 2 proposed that detection of rhythm modulation would be most difficult at extreme initial beat rates, and easiest at moderate initial beat rates (taking 100 bpm as the mid-range value). Since the value of 100 bpm lies roughly equidistantly between 60 and 150 on a logarithmic scale, one would expect these to be approximately equally difficult in the rhythm-modulation conditions. However, rhythm modulation at 150 bpm was statistically significantly more difficult to detect than rhythm modulation at 60 bpm ($t = 6.86$, 446 df, $p < .01$). Hypothesis 2 also proposed that for tempo modulation, detection should get easier the faster the initial beat rate. This is only weakly supported by the cell means in Table XVIII, which display a marginal decrease in beats needed to respond from 60 bpm to 150 bpm, however this difference is not statistically significant ($t = 1.42$, 446 df, $p > .01$).

Hypothesis 2 receives only meagre support from the identification data. The overall interaction between TYPE and INITIAL BEAT RATE is not statistically significant for identification data (for Partial and Marginal Chi-squares, $p > .1$), and it is therefore not listed on Table XIV. However, comparison of individual cell contingencies shown in Table XIX is a little more promising (for this table cell-comparison statistics were computed only between INITIAL-BEAT-RATE levels).

TABLE XIX

EXPERIMENT 2. IDENTIFICATION DATA. CONTINGENCY TABLE FOR 2-WAY INTERACTION. TYPE BY INITIAL BEAT RATE

TYPE	INITIAL BEAT RATE	Wrong	Right	Proportion Correct
Rhythm	60	21	203	.906 (224)
	150	8	216	.964 (224)
Tempo	60	74	150	.670 (224)
	150	25	199	.888 (224)

Separate Chi-square tests on rhythm and tempo modulation reveal the following: the difference between *rhythm-60* and *rhythm-150* is marginal and statistically non-significant (Chi-square = 5.31, 1 df, $p > .01$). On the other hand, the difference between *tempo-60* and *tempo-150* is statistically significant (Chi-square = 29.87, 1 df, $p < .01$) with *tempo-60* being much more difficult than *tempo-150*.

Hypothesis 3 predicts a statistically significant interaction between TYPE and LOCATION. Rhythm modulation should be easier to detect on *beat 2* compared to *beat 1*, whereas there should be no difference between the metric locations of the FAO for tempo modulation. As can be seen from both the beats-to-respond data (Table XII) and the identification data (Table XIV) there was no statistically significant interaction between TYPE and LOCATION: ($F = 1.78$, 1, 27 df, $p > .1$) for beats-to-respond data, (for both Partial and Marginal Chi-Square, $p > .1$) for identification data. Tables XX and XXI below give the cell means and the cell contingencies respectively for the TYPE by LOCATION interaction. For Tables XX and XXI, cell-comparison statistics were computed between LOCATION levels only.

TABLE XX

EXPERIMENT 2, BEATS-TO-RESPOND DATA,
CELL MEANS OF 2-WAY INTERACTION,
TYPE BY LOCATION

TYPE	LOCATION	
	Beat1	Beat2
Rhythm	11.83 (224)	11.17 (224)
Tempo	13.40 (224)	13.27 (224)

TABLE XXI

EXPERIMENT 2, IDENTIFICATION DATA,
CONTINGENCY TABLE FOR 2-WAY INTERACTION, TYPE BY LOCATION

TYPE	LOCATION	Wrong	Right	Proportion Correct
Rhythm	Beat 1	20	204	.911 (224)
	Beat 2	9	215	.960 (224)
Tempo	Beat 1	56	168	.750 (224)
	Beat 2	43	181	.808 (224)

For Table XX, there was no statistically significant difference between *rhythm-early* and *rhythm-late* (t = 1.48, 1 df, p > .01), nor was there between *tempo-early* and *tempo-late* (t = .2, 1 df, p > .01). Likewise, for Table XXI, there was no statistically significant difference between *rhythm-early* and *rhythm-late* (Chi-square = 1.87, 1 df, p > .01) nor was there between *tempo-early* and *tempo-late* (Chi-square = 3.69, 1 df, p > .01).

For the beats-to-respond data alone, there was an unpredicted yet statistically significant two-way interaction between DIRECTION and LOCATION (F = 20.33, 1, 27 df, p < .0005, from Table XII). Table XXII shows the cell means for this interaction; for this table cell-comparison statistics were computed between levels of DIRECTION as well as levels of LOCATION. Table XXII is rendered as a graph in Figure 5.9.

TABLE XXII

EXPERIMENT 2, BEATS-TO-RESPOND DATA, CELL MEANS OF 2-WAY INTERACTION, DIRECTION BY LOCATION

	LOCATION	
DIRECTION	Beat1	Beat2
Early	15.05 (224)	13.67 (224)
Late	10.17 (224)	10.77 (224)

Table XXII reveals that although the difference between *early-beat1* and *early-beat2* did not pass the .01 statistical significance level (t = 2.48, 446 df, p > .01) it was very near the alpha-level (p = .013). On the other hand, there was no statistically significant difference between *late-beat1* and *late-beat2* (t = 1.22, 446 df, p > .01). Since statistically significant differences are found between the

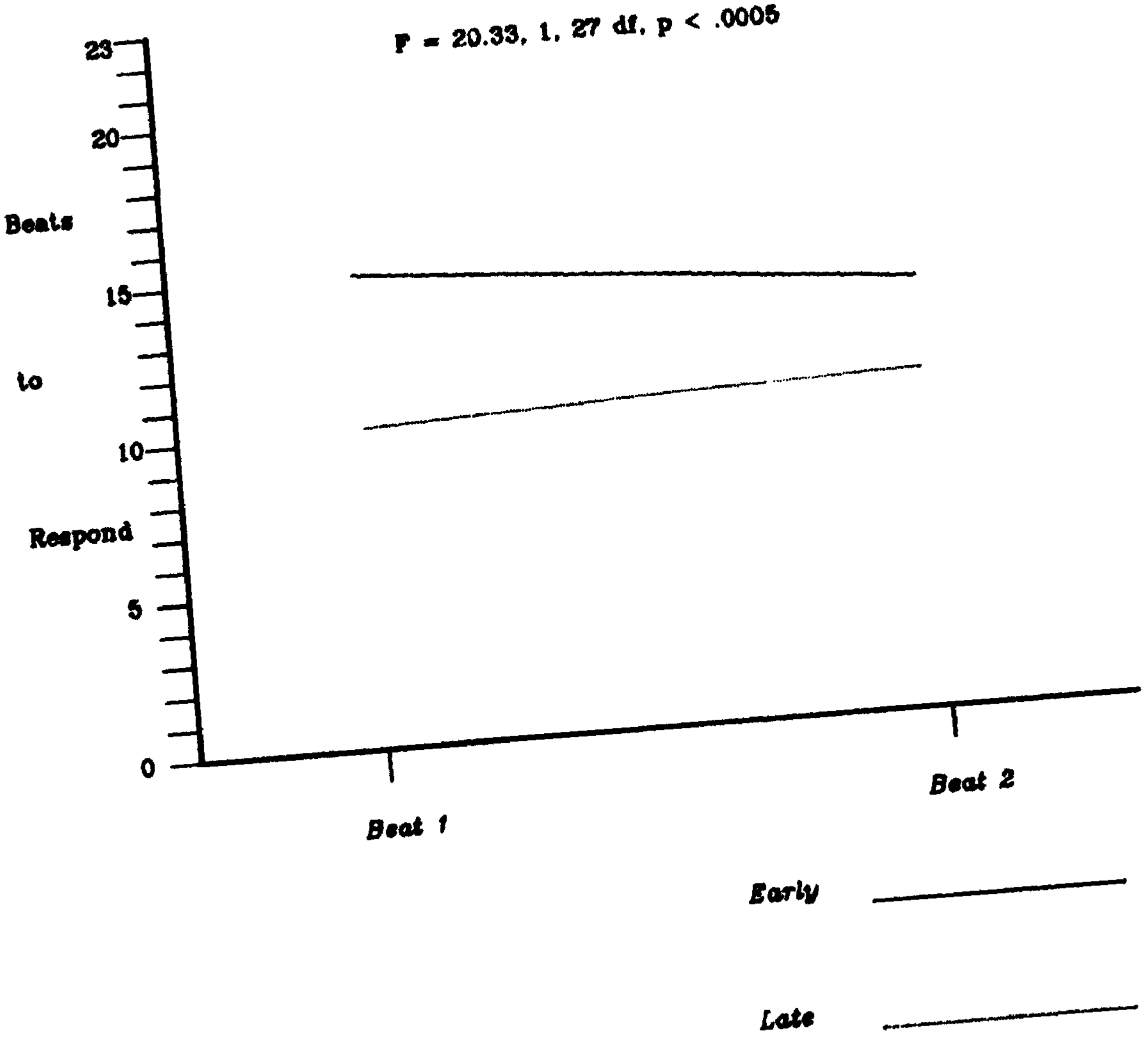


Figure 5.9. Graph of DIRECTION by LOCATION interaction for beats-to-respond data in Experiment 2. The two levels of LOCATION are shown on the x-axis.

pairs *early-beat1/late-beat1* ($t = 9.52$, 446 df, $p < .01$) and *early-beat2/late-beat2* ($t = 5.43$, 446 df, $p < .01$), the overall statistical significance of the DIRECTION by LOCATION interaction must, therefore, lie in the fact that there is a marginal difference between *early-beat1* and *early-beat1* whilst there is no difference between *late-beat1* and *late-beat2*. In other words, the metric location of the FAO was only important when the FAO occurred *early*. One explanation may be that an *early* anomalous onset on beat 2 (whether in rhythm or tempo modulation) would, in effect, shorten the time interval of beat 1. It may be the case that since *beat 1* is normally lengthened in musical performance, a shortening of the same would sound particularly odd and therefore be all the more noticeable.

Finally with regard to Experiment 2, several other check statistics should be mentioned. In Experiment 2, for beats-to-respond data there was no effect owing to which of the two random orders the subjects heard ($t = .16$, 894 df, $p > .75$), nor to starting point of the modulation within the trial ($t = 1.31$, 894 df, $p > .15$). In addition, for identification data, there was no effect owing to order (chi-square $= .3002$, 1 df, $p > .5$) nor to start (chi-square $= 1.54036$, 1df, $p = .2146$). Again, these results are not surprising since the ordering of conditions and the placement of the FAO (bar 2 versus bar 3) were randomly determined.

The correlation coefficient for beats-to-respond data between *original* and *replication* versions of the same condition for each subject was $r = .56$ ($N = 448$, $p < .005$). Again, this positive value is what one would expect, however, if $r = .56$, $r^2 = .32$, indicating that the *original* and *replication* versions only accounted for approximately 32% of each other's variance. The correlation coefficient for identification data between *original* and *replication* versions of the same condition for each subject was phi = .36 (N = 448, p < .005). Again, this positive value is generally what one would predict, however, r^2 = .13, such that, on identification data, the two versions of each stimulus accounted for a meagre 13% of one another's variance. As in Experiment 1, interactions between VERSION and other factors may explain these low r^2 values.

The correlation coefficient between the two dependent variables was again in the direction expected (wrong answers correlating with longer beats-to-respond values) with r = .43 (N = 896, p < .005). However, r^2 = .19, signifying that one dependent variable only accounted for approximately 19% of the variance in the other dependent variable.

To conclude this description of the results for Experiment 2, one can summarize the status of the null hypotheses at this point.

- *Null Hypothesis 1. There will be no statistically significant interaction between TYPE and DIRECTION. Rhythm and tempo modulation will exhibit the same trends* vis à vis *early and late anomalous attacks.* According to the data from Experiment 2, Null Hypothesis 1 can be rejected. Both beats-to-respond data and identification data found a statistically significant interaction between TYPE and DIRECTION following the predicted trends.

- *Null Hypothesis 2. There will be no statistically significant interaction between TYPE and INITIAL BEAT RATE. Rhythm and* *tempo modulation will exhibit essentially the same trends across the musical tempo range.* According to the data from Experiment 2, Null Hypothesis 2 cannot be rejected. Both beats-to-respond and identification data found a statistically significant interaction between TYPE and INITIAL BEAT RATE. However, the precise trends predicted were not supported, and both types of data gave different trends across the initial-beat-rate range. These inconsistencies will be discussed further in Chapter 6.

- *Null Hypothesis 3. There will be no statistically significant interaction between TYPE and LOCATION. Rhythm and tempo modulation will exhibit the same trends across metric location of anomalous onsets.* According to the data from Experiment 2, Null Hypothesis 3 cannot be rejected. Neither beats-to-respond nor identification data found a statistically significant interaction between TYPE and LOCATION. However, an unpredicted yet statistically significant interaction between DIRECTION and LOCATION suggests that the position in the bar of an anomalous onset may play some hitherto unsuspected role in modulation detection. Again, this possibility will be examined in Chapter 6.

Results for Experiment 3

Recall that Experiment 3 was intended primarily to test Hypotheses 1 and 4:

- Hypothesis 1. There will be a statistically significant interaction between TYPE and DIRECTION. *Rhythm* modulation will be indifferent to direction of modulation (*early* versus *late*); *tempo* modulation will be more difficult with *early* than *late*.
- Null Hypothesis 1. There will be no statistically significant interaction between TYPE and DIRECTION. *Rhythm* and *tempo* modulation will exhibit the same trends *vis à vis early* and *late* anomalous attacks.
- Hypothesis 4: There will be a statistically significant effect owing to PIECE in the rhythm-modulation conditions. Owing to the lack of beat subdivision, rhythm modulations in the *Author* piece will be more difficult to detect than equivalent changes in the three pieces with beat subdivision. It is unknown how PIECE will affect tempo modulation perception.
- Null Hypothesis 4: There will be no statistically significant effect owing to PIECE for rhythm modulation. Beat subdivision will not affect detection of rhythm modulation.

The results of the ANOVA performed on beats-to-respond data for Experiment 3 are summarized in Table XXIII. The independent variables are abbreviated as follows: TYP - TYPE, DIR - DIRECTION, PCE - PIECE and VER - VERSION. The same format is used in Table XXIII, as was used in Table III.

TABLE XXIII

EXPERIMENT 3. BEATS-TO-RESPOND DATA. ANALYSIS OF VARIANCE

Source of Variance	Sum of Squares	Df	Mean Square	F	p
TYP	95.67	1	95.67	1.95	.174
Error	1322.07	27	48.97		
DIR	1856.91	1	1856.91	86.64	.000
Error	578.68	27	21.43		
PCE	1119.27	3	373.09	19.39	.000
Error	1558.68	81	19.24		
VER	677.63	1	677.63	44.96	.000
Error	406.93	27	15.07		
TYP:DIR	22.59	1	22.59	.79	.381
Error	768.43	27	28.59		
TYP:PCE	45.59	3	15.20	1.23	.305
Error	1002.02	81	12.37		
TYP:VER	25.05	1	25.05	1.21	.280
Error	557.45	27	20.65		
DIR:PCE	496.14	3	165.38	7.75	.000
Error	1728.86	81	21.34		
DIR:VER	130.11	1	130.11	12.81	.001
Error	274.31	27	10.16		
PCE:VER	35.66	3	11.89	1.14	.339
Error	847.33	81	10.46		
TYP:DIR:PCE	117.37	3	39.13	2.58	.059
Error	1229.80	81	15.18		
TYP:DIR:VER	22.46	1	22.46	2.69	.113
Error	225.35	27	8.35		
TYP:PCE:VER	44.39	3	14.80	1.30	.279
Error	918.50	81	11.34		
DIR:PCE:VER	356.03	3	118.68	9.91	.000
Error	970.13	81	11.98		
TYP:DIR:PCE:VER	82.03	3	27.34	2.04	.115
Error	1085.93	81	13.41		
Explained	5126.90	31	165.38	8.00	.000
Residual	17872.71	864	20.69		
Total	22999.61	895	25.70		

The results of the Survival Analysis are given in Table XXIV, which uses the same abbreviations as Table XXIII. As can be seen from comparing Tables XXIII and XXIV, the same main effects were found to be statistically significant using both ANOVA and Survival Analysis (that is, all effects except TYPE). As in Experiments 1 and 2, such an agreement between the two tests suggests that assigning the beats-to-respond value of 23 to missing responses was justified.[3]

TABLE XXIV

EXPERIMENT 3, BEATS-TO-RESPOND DATA, SURVIVAL ANALYSIS

Source of Variance	p
TYP	> .0100
DIR	< .0001
PCE	< .0001
VER	< .0001

---

[3] In order to test the data-conversion null hypothesis for Experiment 3, an ANOVA was also performed on the raw reaction-time data (in milliseconds). This analysis found that, generally, those main main effects and interactions that were statistically significant for raw reaction-time data were the same as those that were statisticallly significant for beats-to-respond data. There was one exception. The main effect TYPE was statistically significant ($F = 7.71$, 1, 27 df, $p < .01$) for raw reaction-time data when it was not statistically significant for beats-to-respond data. Since this result outstripped the significance level of the corresponding main effect for beats-to-respond data, the data-conversion null hypothesis could be rejected in all cases for Experiment 3.

Table XXV shows the overall results of the Log-Linear analysis of the identification data. Only those results that approach statistical significance (that is, $p < .1$) are shown on this Table. Again, IDN stands for identification data.

TABLE XXV

EXPERIMENT 3, IDENTIFICATION DATA, LOG-LINEAR ANALYSIS

Dependent Variable	Independent Variable(s)	Df	Partial Chi-Squared	p	Marginal Chi-Squared	p
IDN:	VER	1	5.959	.015	5.942	.015
IDN:	TYP:DIR:PCE	3	9.899	.019	11.260	.010

In Experiment 3, the hypotheses concerned themselves with the interaction between TYPE and DIRECTION and the main effect of PIECE (or possibly the interaction between TYPE and PIECE). Surprisingly, contrary to Hypothesis 1 and to the findings of the other two experiments, there was no statistically significant interaction between TYPE and DIRECTION for beats-to respond data ($F = .79$, 1, 27 df, $p > .25$). Table XXVI gives the cell means for the TYPE by DIRECTION interaction; cell-comparison statistics were computed between levels of DIRECTION only.

TABLE XXVI

EXPERIMENT 3. BEATS-TO-RESPOND DATA.
CELL MEANS OF 2-WAY INTERACTION.
TYPE BY DIRECTION

	DIRECTION	
TYPE	Early	Late
Rhythm	10.68 (224)	8.12 (224)
Tempo	10.34 (224)	7.15 (224)

As can be seen from Table XXVI, the cell means for this interaction reveal that although tempo modulation followed the previous experiments in that *tempo-early* was statistically significantly more difficult than *tempo-late* (t = 6.77, 446 df, p < .01), rhythm modulation produced unexpected results. Essentially, the rhythm-modulation cell means followed the same trend as tempo-modulation, with *rhythm-late* being statistically significantly easier than *rhythm-early* (t = 5.76, 446 df, p < .01). In other words, subjects seemed to be responding to rhythm modulation in the same way that they responded to tempo modulation vis a vis DIRECTION. This unexpected trend will be discussed further in Chapter 6.

From Tables XXV above and XXVII below it can be seen that the TYPE by DIRECTION interaction is also statistically non-significant for the identification data (p > .1 for Partial and Marginal Chi-Squares). For Table XXVII, cell-comparison statistics were computed between levels of DIRECTION only.

TABLE XXVII

EXPERIMENT 3, IDENTIFICATION DATA, CONTINGENCY TABLE FOR 2-WAY INTERACTION, TYPE BY DIRECTION

TYPE	DIRECTION	Wrong	Right	Proportion Correct
Rhythm	Early	18	206	.920 (224)
	Late	15	209	.933 (224)
Tempo	Early	11	213	.951 (224)
	Late	14	210	.938 (224)

In the case of Table XXVII, there is simply no statistically significant difference either between *rhythm-early* and *rhythm-late*, or *tempo-early* and *tempo-late*. Generally, in fact, the identification data in Experiment 3 turned up hardly any statistically significant results. There was a ceiling effect; subjects seem to have found the identification task very easy with the result that there were very few wrong answers. This reduced any possible contrasts between cell contingencies.

Overall, in support of Hypothesis 4, the main effect PIECE was highly significant on the beats-to-respond data ($F = 19.39$, 3, 81 df, $p < .0005$, from Table XXIII). The cell means for this effect together with individual cell comparisons are given again as Table XXVIII below, and as a graph in Figure 5.10.

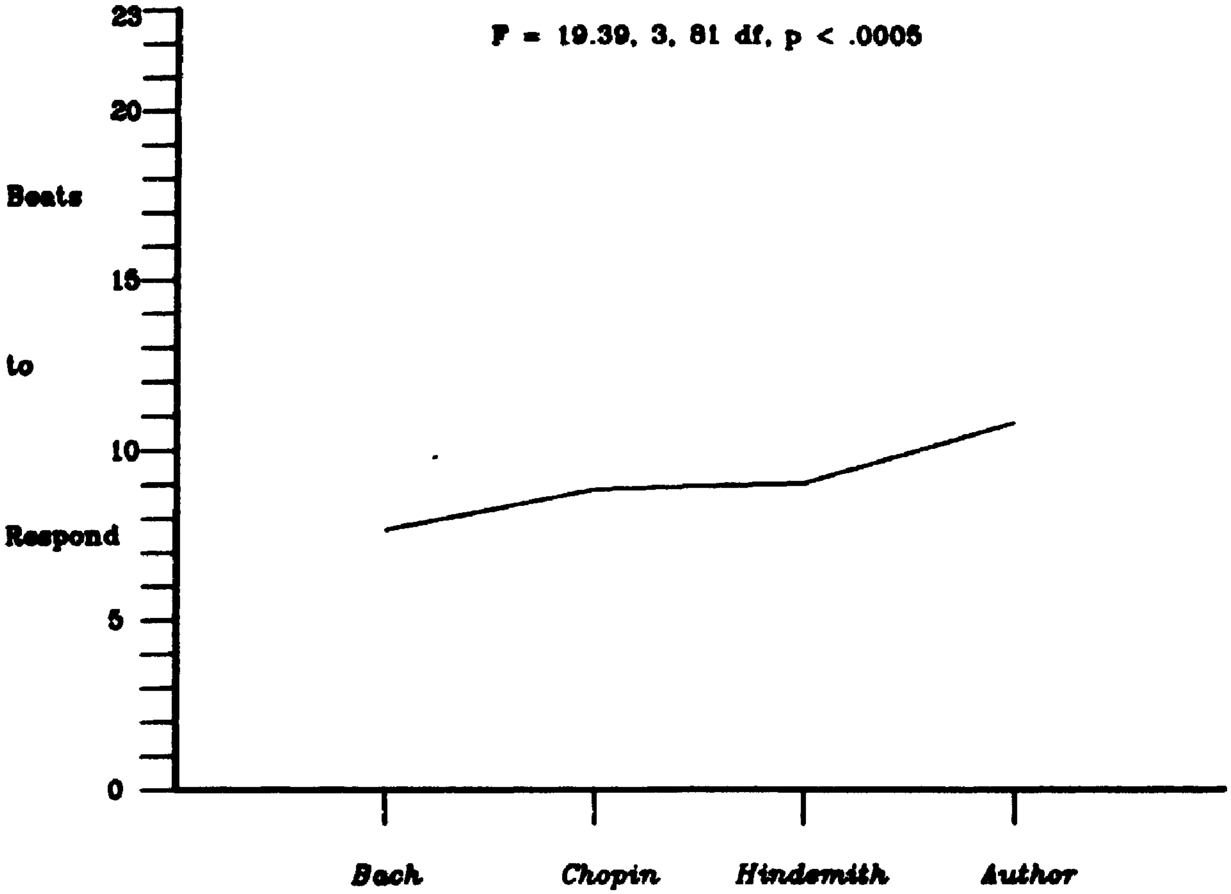


Figure 5.10. Graph of main effect PIECE for beats-to-respond data in Experiment 3. The four levels of PIECE are shown on the x-axis.

TABLE XXVIII

EXPERIMENT 3, BEATS-TO-RESPOND DATA,
CELL MEANS OF MAIN EFFECT, PIECE

PIECE			
Bach	Chopin	Hindemith	Author
7.66 (224)	8.84 (224)	9.00 (224)	10.79 (224)

As in previous interaction tables with more than two cells on any variable, those cells joined by a line are not statistically significantly different at the .01 level according to Duncan's multiple-range test. Perusal of Table XXVIII reveals that rhythm and tempo modulation were detected with the following decreasing order of difficulty across the pieces: *Author*, *Hindemith*, *Chopin* and *Bach*. It would be plausible to attribute this statistically significant difference entirely to familiarity of musical style (with *Bach* being the best known) were it not for the fact that the *Author* piece was specially composed in a Common-Practice tonal style reminiscent of the *Bach*. Therefore, the difference between the *Author* piece and the other three must lie in some factor other than musical style. By far the most probable candidate is the lack of beat subdivision in the *Author* piece. Conversely, in line with Hypothesis 4, sixteenth-note beat subdivision in the *Bach*, *Chopin* and *Hindemith* seems to have rendered temporal change detection easier. As can be surmised from Table XXV, there was no statistically significant effect owing to PIECE on the identification data ($p > .1$).

Finally with regard to Hypothesis 4, it will be remembered that the possibility of an interaction between TYPE and PIECE was also entertained. However, this two-way interaction was not statistically significant for either the beats-to-respond data (from Table XXIII, F = 1.23, 3, 81 df, p > .25) nor for the identification data (p > .1, from Table XXV). Given this lack of overall statistical significance on the two types of dependent variable, no cell-comparison statistics were performed. However, for reference, the cell means and cell contingencies for the TYPE by PIECE interaction are given below in Tables XXIX and XXX.

TABLE XXIX

EXPERIMENT 3, BEATS-TO-RESPOND DATA,
CELL MEANS OF 2-WAY INTERACTION, TYPE BY PIECE

	PIECE			
TYPE	Bach	Chopin	Hindemith	Author
Rhythm	8.34 (112)	9.11 (112)	9.31 (112)	10.84 (112)
Tempo	6.98 (112)	8.57 (112)	8 68 (112)	10.74 (112)

TABLE XXX

EXPERIMENT 3, IDENTIFICATION DATA, CONTINGENCY TABLE FOR 2-WAY INTERACTION, TYPE BY PIECE

	TYPE					
	Rhythm			Tempo		
PIECE	Wrong	Right	Proportion Correct	Wrong	Right	Proportion Correct
Bach	7	105	.938 (112)	3	109	.973 (112)
Chopin	5	107	.955 (112)	10	102	.911 (112)
Hindemith	11	101	.902 (112)	5	107	.955 (112)
Author	10	102	.911 (112)	7	105	.938 (112)

The above interactions and main effects deal with the hypotheses for Experiment 3. Two other related, unpredicted, yet statistically significant interactions deserve mention as well. According to Table XXIII, DIRECTION by PIECE was statistically significant on the beats-to-respond data only (F = 7.75, 3, 81 df, p < .0005). Table XXXI gives the cell means for this interaction; for this table cell-comparison statistics are computed between levels of PIECE only.

TABLE XXXI

EXPERIMENT 3, BEATS-TO-RESPOND DATA,
CELL MEANS OF 2-WAY INTERACTION, DIRECTION BY PIECE

	PIECE			
DIRECTION	Bach	Chopin	Hindemith	Author
Early	8.25 (112)	10.57 (112)	9.93 (112)	13.29 (112)
Late	7.07 (112)	7.11 (112)	8.07 (112)	8.28 (112)

Glancing at the cell means in Table XXXI, it can be seen that an *early* FAO in the *Chopin* piece was statistically significantly more difficult to perceive than in the *Bach* piece (Duncan $p < .01$). In addition, an *early* FAO in the *Author* piece was statistically significantly more difficult than all three other pieces. In contrast, for a *late* FAO there were no statistically significant cell-mean comparisons (for all Duncan $p > .01$). This interaction is difficult to explain and may have to do with other characteristics of the pieces not already accounted for such as, for instance, the size and direction of melodic motion at the time in the piece when the anomalous onsets were presented. This and other extraneous sources of variance will be evaluated in Chapter 6.

For the sake of completeness, the 2-way interaction DIRECTION by PIECE must be considered in the light of the statistically significant three-way interaction DIRECTION by PIECE by VERSION for beats-to-respond data only ($F = 9.91$, 3, 81 df, $p < .0005$ according to Table XXIII). The cell means for the DIRECTION by PIECE by VERSION interaction are given below in Table XXXII. For Table XXXII, cell-comparison statistics were computed between levels of PIECE and VERSION but not between levels of DIRECTION.

TABLE XXXII

EXPERIMENT 3, BEATS-TO-RESPOND DATA,
CELL MEANS OF 3-WAY INTERACTION, DIRECTION BY PIECE BY VERSION

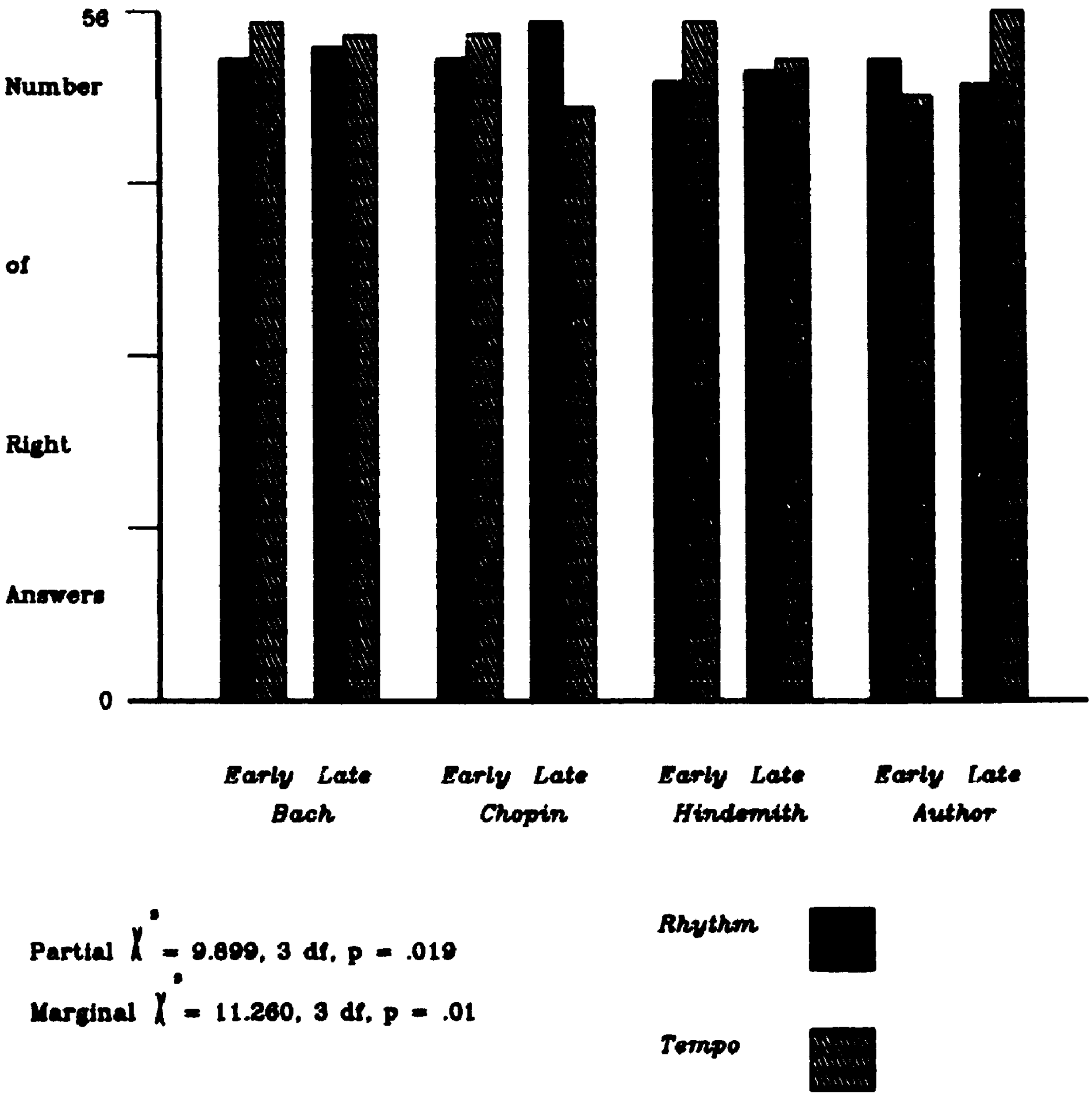
		PIECE			
DIRECTION	VERSION	Bach	Chopin	Hindemith	Author
Early	Original	10.38 (56)	11.19 (56)	10.27 (56)	15.21 (56)
	Replication	6.12 (56)	9.95 (56)	9.59 (56)	11.38 (56)
Late	Original	6.97 (56)	8.32 (56)	8.79 (56)	8.39 (56)
	Replication	7.17 (56)	5.89 (56)	7.35 (56)	8.16 (56)

Table XXXII reveals that for the *late-original* and *late-replication* conditions there were no statistically significant cell differences across the four pieces (Duncan $p > .01$ for all comparisons). However, for *early-original* the *early-Author-original* cell was statistically significantly more difficult than the other three pieces (Duncan $p < .01$). Also, for *early-replication* the *early-Bach-*

replication cell was statistically significantly easier than the other three pieces (Duncan p < .01).

Looking at Table XXXII a second way, three DIRECTION-PIECE conditions exhibited statistically significant learning between *original* and *replication* versions: the *Bach* piece with an *early* FAO (t = 5.21, 110 df, p < .01), the *Author* piece with an *early* FAO (t = 4.29, 110 df, p < .01), and the *Chopin* piece with a *late* FAO (t = 3.08, 110 df, p < .01). All other comparisons between *original* and *replication* versions for DIRECTION-PIECE conditions were statistically non-significant (t < 2.36, 110 df, p > .01). Again, this highly specific interaction is difficult to interpret and may arise from unaccounted idiosyncrasies of the pieces in question.

Finally, with regard to the main statistical tests in Experiment 3, further tentative support for the idea that beat subdivision affects the way in which temporal change is perceived may be found from the identification data. Table XXV shows that the 3-way interaction TYPE by DIRECTION by PIECE is statistically significant for one of the measures in that table (Marginal Chi-square = 11.26, 3 df, p = .01) and approaching statistical significance for the other measure (Partial Chi-square = 9.90, 3 df, p = .019). Table XXXIII shows the contingency cells organized by PIECE and TYPE comparing directions for this interaction. Figure 5.11 reproduces Table XXXIII as a histogram.



Partial χ^2 = 9.899, 3 df, p = .019

Marginal χ^2 = 11.260, 3 df, p = .01

Figure 5.11. Histogram of TYPE by DIRECTION by PIECE interaction for identification data in Experiment 3. The two levels of DIRECTION in combination with the four levels of PIECE are shown on the x-axis. The total number of responses per cell is shown atop the y-axis.

TABLE XXXIII

EXPERIMENT 3, IDENTIFICATION DATA,
CONTINGENCY TABLE FOR 3-WAY INTERACTION,
TYPE BY DIRECTION BY PIECE

		DIRECTION					
		Early			Late		
PIECE	TYPE	Wrong	Right	Proportion Correct	Wrong	Right	Proportion Correct
Bach	Rhythm	4	52	.929 (56)	3	53	.946 (56)
	Tempo	1	55	.982 (56)	2	54	.964 (56)
Chopin	Rhythm	4	52	.929 (56)	1	55	.982 (56)
	Tempo	2	54	.964 (56)	8	48	.857 (56)
Hindemith	Rhythm	6	50	.893 (56)	5	51	.911 (56)
	Tempo	1	55	.982 (56)	4	52	.929 (56)
Author	Rhythm	4	52	.929 (56)	6	50	.893 (56)
	Tempo	7	49	.875 (56)	0	56	1.000 (56)

For the contingencies in Table XXXIII it was not possible to calculate individual cell-comparison Chi-square statistics since 50% of the expected frequencies in cells are less than 5. However, it is apparent from Figure 5.11 that tempo modulation in the *Author* piece reveals a different trend across DIRECTION conditions compared with all the other pieces: for the *Author* piece *tempo-early* was more difficult than *tempo-late*, whereas for the *Bach*, *Chopin* and *Hindemith* pieces *tempo-late* was more difficult than *tempo-early*. For rhythm modulation, there was no systematic trend. However, as in all other identification data for Experiment 3, the number of wrong answers was so small that any conclusion drawn from these data is at best tentative.

Finally with regard to Experiment 3, several other check statistics should be mentioned. In Experiment 3, for beats-to-respond data, there was no effect owing to trial order ($t = .53$, 894 df, $p > .5$), nor to the starting point of modulation within each trial ($t = 1.29$, 894 df, $p > .15$). The same was true for identification data with both order (chi-square = .698, 1 df, $p > .4$) and start of modulation (bar 2 versus bar 3) (chi-square = .461, 1df, $p > .4$).

The correlation coefficient for beats-to-respond data between *original* and *replication* versions of the same condition for each subject was $r = .46$ ($N = 448$, $p < .005$). It therefore follows that $r^2 = .22$ indicating that the *original* and *replication* versions only accounted for approximately 22% of each others variance. The correlation coefficient for identification data between *original* and *replication* versions of the same condition for each subject was phi = .25 ($N = 448$, $p < .005$). In this case $r^2 = .06$ indicating that *original* and *replication* versions only accounted for 6% of each other's variance. The correlation coefficient between the two dependent variables was again in the direction expected (wrong answers correlating with longer beats-to-respond values) with $r = .26$ ($N = 896$, $p < .005$). However, $r^2 = .07$, signifying that one dependent variable only accounted for approximately 7% of the variance in the other dependent variable. Again, these low r^2 values will be discussed in Chapter 6.

To conclude this description of the results of Experiment 3, one can summarize the status of the null hypotheses at this point:

- *Null Hypothesis 1. There will be no statistically significant interaction between TYPE and DIRECTION. Rhythm and tempo modulation will exhibit the same trends vis a vis early and late anomalous attacks.* According to the data from Experiment 3, Null Hypothesis 1 cannot be rejected. There was no statistically significant interaction between TYPE and DIRECTION.

- *Null Hypothesis 4. There will be no statistically significant effect owing to PIECE for rhythm modulation. Beat subdivision will not affect detection of rhythm modulation.* According to the data from Experiment 3, Null Hypothesis 4 can be rejected. There was a statistically significant main effect owing to PIECE, which was due to the *Author* piece making modulation detection statistically significantly more difficult than modulation detection in other pieces.

Conclusion

The results may be summarized in terms of the four experimental hypotheses as they were outlined at the beginning of this chapter:

- Hypothesis 1. Experiments 1 and 2 found statistically significant interactions between TYPE and DIRECTION supporting Hypothesis 1. However, Experiment 3 found no such interaction, signifying that for this third experiment, Null Hypothesis 1 could not be rejected.

- Hypothesis 2. Experiments 1 and 2 found statistically significant interactions between TYPE and INITIAL BEAT RATE. However, only some of the results matched the specific trends across the musical tempo range predicted by Hypothesis 2. For rhythm modulation, beats-to-respond data evinced a slower optimum INITIAL BEAT RATE than the 100 bpm predicted (75 bpm for Experiment 1 and 60 bpm for Experiment 2). However, for Experiment 1, the optimum tempo for rhythm modulation detection with identification data was 120 bpm, faster than 100 bpm. For tempo modulation, beats-to-respond data generally followed the predicted trend, that is with detection becoming easier the faster the INITIAL BEAT RATE. This trend was matched exactly by the identification data and partially by the beats-to-respond data, the latter evincing a fall in reaction time up to 190 bpm and a sudden rise between 190 and 300 bpm.

- Hypothesis 3. Experiment 2 found no support for Hypothesis 3, namely there was no statistically significant interaction between TYPE and LOCATION. For this reason Null Hypothesis 3 could not be rejected. However, an unpredicted significant interaction between DIRECTION and LOCATION on beats-to-respond data suggested that LOCATION does affect modulation detection in some unforeseen way.

- Hypothesis 4. Experiment 3 found a statistically significant main effect owing to PIECE, which tentatively supports Hypothesis 4 that beat subdivision aids modulation detection. The (uncontrolled) influence of musical style is here discounted since the two pieces that exhibited the greatest mean difference (the *Bach* and *Author* pieces) were in similar harmonic/melodic styles.

The above results, the way they tie in with the established literature, and the way they both support and necessitate revision of the model outlined in Chapter 3, will form the focus of the final chapter in this dissertation, Chapter 6.

CHAPTER 6

DISCUSSION AND CONCLUSION

By way of conclusion, this chapter summarizes the status of the various hypotheses, and interprets these findings in light of both the literature reviewed in Chapter 2 and the perceptual model outlined in Chapter 3. Possible confounds in the experimental procedures are identified and addressed and suggestions for future research are made.

Summary and Discussion of Results

The statuses of the four experimental hypotheses as they were tested in the three experiments for the two types of data are summarized in Figure 6.1. A tick indicates that a specific result supported the hypothesis with statistically significant results in the manner predicted. A tick in parentheses indicates a statistically significant results, but evincing a trend that only partly matches the predicted trend. A zero indicates no statistically significant results and therefore an inablility to reject the null hypothesis.

Hypothesis 1

- *Hypothesis 1. There will be a statistically significant interaction between TYPE and DIRECTION. Rhythm modulation will be*

Hypothesis	Interaction	Experiment	BTR Data	IDN Data
1	TYP:DIR	1	✓	✓
		2	✓	✓
		3	0	0
2	TYP:IBR	1	(✓)	(✓)
		2	(✓)	(✓)
3	TYP:LOC	2	0	0
4	PCE	3	✓	0

Figure 6.1. Summary of status of the four experimental hypotheses.

indifferent to direction of modulation (early versus late): tempo modulation will be more difficult with early than with late.

Hypothesis 1 received strong support from both Experiment 1 and Experiment 2 for both the beats-to-respond data and the identification data. The beats-to-respond data found no significant difference at the .01 level between early and late conditions for rhythm modulation. Of course, this failure to pass the .01 alpha level is not proof that *rhythm-early* and *rhythm-late* are equally difficult to perceive, but it is a failure to prove that there is any difference in the difficulty with which they are detected. Similar trends were recorded for Experiments 1 and 2 for the identification data for rhythm modulation. In this case, the identification data found no or negligible difference between the numbers of wrong answers for *rhythm-early* and *rhythm-late* conditions (and these differences were not statistically significant).

For tempo modulation, the beats-to-respond data for both Experiment 1 and Experiment 2 revealed a statistically significant difference between early and late conditions. In both experiments, *tempo-early* (tempo acceleration) was statistically significantly more difficult than *tempo-late* (tempo deceleration). Likewise, the identification data revealed a statistically significant difference in the numbers of wrong answers between *tempo-early* and *tempo-late*. As was the case for the beats-to-respond data, identification data showed that *tempo-early* (tempo acceleration) was more difficult to detect than *tempo-late* (tempo deceleration).

The one set of recalcitrant results so far as Hypothesis 1 was concerned were garnered from Experiment 3. For the beats-to-respond data there was no statistically significant interaction between TYPE and DIRECTION. Tempo modulation followed the same trend as had been observed in Experiments 1 and 2, viz., a statistically significant difference in mean beats-to-respond between *tempo-early* and *tempo-late*, with *tempo-early* being more difficult than *tempo-late*. However, instead of the roughly equal means for *rhythm-early* and *rhythm-late* that had been observed in Experiments 1 and 2, in Experiment 3, surprisingly, rhythm modulation evinced the same trend between early and late as was observed for tempo modulation. In other words, a statistically significant difference was observed between *rhythm-early* and *rhythm-late*, with early being more difficult than late.

One possible interpretation of the results of Experiment 3 may be framed in terms of the model outlined in Chapter 3. With the extracts of keyboard music used in this experiment, it may be the case that both the SHORT AUDITORY STORE (SAS) and the OSCILLATOR BANK were working in tandem for rhythm-modulation perception. As in the simple drum stimuli, oscillators in the OSCILLATOR BANK would be entrained to the time intervals in the stimulus. However, in Experiment 3, the complexity of the stimuli is considerably greater than those of Experiments 1 and 2. Purely in rhythmic terms, four times as many events are happening in one beat in Experiment 3 as happened in one beat in Experiments 1 and 2. This is without taking into account other aspects of the signal that are present in Experiment 3, but not in Experiments 1 and 2 (melodic pitch contour, voice-leading in other parts, the resulting harmony etc.). For this reason, it is plausible that an SAS buffer would fill up more quickly in Experiment 3 than it did in Experiments 1 and 2. If, for argument's sake, the length of one chunk of SAS in Experiment 3 was about one beat, then comparison between adjacent beats, including comparison of temporal properties, could be achieved by comparing the contents of adjacent buffers in SAS. If this were the case, rhythm-modulation perception would be affected by all the inexactitudes of SAS operation that affected only tempo-modulation detection in Experiments 1 and 2. This would explain why rhythm modulation followed the same trend as tempo modulation in Experiment 3. If the above account is true, then some re-interpretation of Hypothesis 4 is also required.

If this explanation sounds reasonable, it should nevertheless be accepted with caution. It is a *post hoc* explanation that is required because the original model did not adequately account for all the results. The failure to predict this specific result in Experiment 3 may arise from the fact that the notion of "stimulus complexity" is not outlined in any detail in the model. It seems intuitively reasonable to assume that extracts of keyboard music are "more complex" than drum-strike stimuli, however, the extent to which they are more complex cannot be quantified in the current model. It follows, therefore, that it is not possible to predict when a stimulus is complex enough to cause the SAS to divide the flow of musical events up into chunks of only one beat in length.

TYPE by DIRECTION in the previous literature. No previous studies have used rhythm modulation in the way it has been used here. For this reason, there are no earlier findings with which the results of the present experiments can be directly compared regarding rhythm modulation. For instance, the literature reviewed in Chapter 2 that adopted the Weber's-law paradigm did not test directly for the effect on duration discrimination of direction (see again Getty 1975; Halpern and Darwin 1982; Hirsh, Monahan, Grant and Singh 1990).

The results of the present experiments so far as tempo modulation was concerned support much of the literature cited in Chapter 2. In all of the present experiments, tempo modulation with a late anomalous onset (tempo deceleration) was found to be easier to perceive than tempo modulation with an early anomalous onset (tempo acceleration) of the same magnitude. This finding bolsters similar results observed by Duke (1994), Kuhn (1974), Madsen (1979), Royal (1993) and Wang (1984). This finding contradicts other results obtained by Geringer and Madsen (1984) and Madsen, Duke and Geringer (1984).

Hypothesis 2

- *Hypothesis 2. There will be a statistically significant interaction between TYPE and INITIAL BEAT RATE. Rhythm modulation will be most difficult to detect at extreme initial beat rates, whereas tempo modulation will be most difficult at slow initial beat rates.*

Rhythm Modulation. Of all the hypotheses advanced in the course of this dissertation, the results of Experiments 1 and 2 speak to Hypothesis 2 in the most contradictory manner. For beats-to-respond data the rhythm-modulation results of both experiments suggest that modulation detection became more difficult the faster the initial beat rate. For Experiment 1 the optimum initial beat rate for rhythm modulation detection was 75 bpm (although there was no statistically significant difference between the mean detection values for 45, 75 and 120 bpm). For Experiment 2 the preferred initial beat rate for rhythm modulation detection was 60 bpm. For both experiments, then, initial beat rates faster than 120 bpm made rhythm modulation detection comparatively difficult according to beats-to-respond data. For the faster initial beat rates, this trend is what one would expect: those initial beat rates faster than the preferred tempo range made rhythm-modulation detection more difficult. However, for the slower initial beat rates, the expected trend did not occur, particularly at the very slow initial beat rate of 45 bpm where rhythm-modulation detection was almost as easy as that at 75 bpm.

For the identification data for Experiments 1 and 2, one finds a different trend for rhythm modulation, which again does not exactly match the one predicted. With these data, for Experiment 1, one finds the highest numbers of wrong answers at the extreme initial beat rates of 45 and 300 bpm (as predicted). However, the optimum range (that with the lowest numbers of wrong answers) lies between 120 and 190 bpm, which is faster than the preferred tempo range identified in the model in Chapter 3. The findings of Experiment 1 in this regard are supported by those of Experiment 2: for identification data, rhythm modulation at 150 bpm elicited statistically significantly fewer wrong answers than the same at 60 bpm.

For rhythm modulation alone, then, the beats-to-respond data suggest a preferred tempo range from 45 to 120 bpm whilst the identification data indicate one from 120 to 190 bpm. The fact that beats-to-respond data favoured slow to medium initial beat rates and that identification data favoured medium to fast initial beat rates is somewhat contradictory. One possible explanation is that at least some of the subjects mentally subdivided the initial beat rate of 45 bpm into twos. This action would effectively activate oscillators at 90 bpm in the listener's brain as well as those entrained to the 45-bpm stimulus. If the subject did this, the relative asynchrony of any rhythmic anomaly at 45 bpm would be doubled at 90 bpm, making detection twice as easy. Two pieces of evidence suggest that the strategy of mental subdivision was adopted by some the subjects: (1) in the informal debriefing after individual sessions for Experiment 1, several subjects intimated that they had had to subdivide the very slow stimuli, either by tapping or in their heads; (2) the standard deviation for the 45-bpm condition is larger than all the other initial-beat-rate conditions. This high standard-deviation value is due to some extreme values (both low and high) in the data set for that condition, suggesting the use of more than one listening strategy. Naturally, if mental subdivision were found to be a useful strategy with very slow stimuli, a subject would try this regardless of the type of modulation that occurred. However, the advantage of mental subdivision would only be reflected in the beats-to-respond data for rhythm modulation, since, according to the model at least, tempo modulation detection would be mediated by the SAS rather than the OSCILLATOR BANK.

Of course, both of the types of evidence for mental subdivision of very slow beats mentioned above are informal and circumstantial. With the present experimental set-up, there was no way of knowing exactly to what extent subjects mentally subdivided the 45-bpm stimuli.

However, further evidence that this type of subdivision is a common way of handling very slow stimuli may be found in the previous experimental literature (see again Duke 1989b; Duke, Geringer and Madsen 1991; Geringer, Duke and Madsen 1992).

An alternative explanation, which also elucidates to some extent the different trends revealed by the two data types, is that detecting an anomalous IOI and classifying that anomaly require mental processes that work over different time frames. To detect a rhythm modulation, a subject need only compare the time intervals of two adjacent beats. This process, even at an initial beat rate of 45 bpm, would entail integrating a total time span of roughly 2.7 seconds. However, to classify the type of modulation (i.e., to ascertain that the modulation was not in fact a tempo modulation), one would need to hear at least one bar, entailing integration of a time span of at least 5.3 seconds. This may illuminate why, for rhythm modulation, beats-to-respond data favoured slow-to-medium initial beat rates, whereas identification data favoured medium-to-fast initial beat rates where the to-be-integrated time interval would be shorter.

For both sets of data, performance at an initial beat rate of 300 bpm deteriorated, suggesting that this tempo is too fast for accurate detection of rhythm modulation whichever way one measures detection performance. If one accepts the model outlined in Chapter 3, this consistency between the two types of data suggests that there is a rapid decline in the OSCILLATOR BANK's capability to entrain to stimuli between 190 and 300 bpm.

Tempo modulation. For tempo modulation, beats-to-respond data revealed a consistent trend of decreasing difficulty from 45 bpm to 190 bpm in Experiment 1. This trend is matched in Experiment 2 with a decline in difficulty being observed from 60 to 150 bpm. Although the result in Experiment 2 was not found to be statistically significant according to cell-comparison statistics, this lack of statistical significance may be due to the fact that 60 and 150 bpm are not far enough apart on the initial-beat-rate scale to register a great enough mean difference. After all, in Experiment 1, of all the initial-beat-rate conditions for tempo modulation, only the *tempo-45* and the *tempo-190* cell means were statistically significantly different from each other according to Duncan's multiple range test.

Between *tempo-190* and *tempo-300* in Experiment 1, there was an up-turn in difficulty for beats-to-respond data. Again, this result suggests that there is a threshold at some point between 190 and 300 bpm where tempo modulation detection begins to get more difficult rather than easier. Although one should be cautious about combining data from earlier studies (and therefore from different experimental approaches), evidence from previous literature can throw further light on where exactly this change in trend might begin. Recall that Ellis (1991) found an overall decrease in the mean threshold difference for tempo modulation detection between 192 and 228 bpm. Similarly, Royal (1993; see Appendix A) found that tempo modulation at 245 bpm was significantly more "obvious" than modulation at 105 bpm. Therefore, it may be conjectured that the up-turn in difficulty of perceiving tempo modulation occurs between 245 and 300 bpm. In other words, the up-turn occurs when successive beats occur more than about 4 times a second, or are separated by less than about 245 milliseconds.

The aforementioned up-turn in difficulty at rates faster than four events per second is contrary to the model outlined in Chapter 3, which predicted easier detection of tempo modulation the faster the tempo. However, in general, beats-to-respond data supports the model in both Experiments 1 and 2.

Further support for Hypothesis 2 so far as tempo modulation perception is concerned, is gathered from the identification data for Experiments 1 and 2. Experiment 1 found a consistent decrease in the number of wrong answers for tempo modulation, the faster the initial beat rate. Likewise, Experiment 2 found a decrease in the numbers of wrong answers from 60 to 150 bpm. Note that for this type of data, Experiment 1 found *no* up-turn in difficulty between 190 and 300 bpm. Again, the reason for this difference between beats-to-respond and identification data may lie in the fact that the more rapidly the stimulus was completed, the shorter the time span that subjects were required to integrate in order to classify a modulation type, and the easier that classification task would be.

TYPE by INITIAL BEAT RATE in previous literature. Again, as was the case with Hypothesis 1, since no previous literature exists that employed rhythm modulation in the same way as these studies, there is no body of experimentation that can be used for purposes of direct comparison. However, it is perhaps worth recalling that literature on duration discrimination using the Weber's-law paradigm found an increase in the Weber fraction for base durations outside the range 200 to 2000 milliseconds (in terms of metronome markings, 300 to 30 bpm). Note that an inter-beat interval (IBI) for a stimulus with an initial beat rate of 45 bpm would be within this range. However, an IBI for a stimulus with an initial beat rate of 300 bpm would be at the extreme end of this range. This may throw further light on the fact that, at 45 bpm, rhythm-modulation detection was reasonably easy, but much more difficult at 300 bpm. In other words, the preferred-duration range identified by Weber's law experiments is more in line with the preferred initial beat rate range observed in Experiments 1 and 2, than the range hypothesized by the model outlined in Chapter 3. Of course, Weber's law does not posit a hypothetical mechanism or process by which duration discrimination might take place; Weber's law, after all, is only an idealized mathematical description of observed behaviours. However, these findings do suggest that the central, preferred tempo range of the OSCILLATOR BANK should perhaps be extended to include slower beat-rates such as 45 bpm.

As regards tempo modulation, the greater facility for modulation detection at faster initial beat rates (except 300 bpm for beats-to-respond data) is in line with the results of studies by Ellis (1991; 1992) Wapnick (1980) and Royal (1993), and contrary to findings by Dorhout (1980) and Wang and Salzberg (1984).

Interestingly, there was no interaction between DIRECTION and INITIAL BEAT RATE for tempo modulation. Recall that Brittin (1992; 1993), Ellis (1991) and Yarbrough (1987) found that subjects most accurately identified tempo decelerations for slow pieces, and tempo accelerations for fast pieces. As can be seen from reviewing Table X and Figure 5.5 from the previous chapter, this particular interaction was not replicated in the present studies.

Hypothesis 3

- *Hypothesis 3. There will be a statistically significant interaction between TYPE and LOCATION. Rhythm modulation will be more difficult at beat 1 than at beat 2; tempo modulation will be equally difficult to detect at beats 1 and 2.*

Experiment 2 alone was designed to test Hypothesis 3. On both beats-to-respond and identification data there was no statistically significant interaction between TYPE and LOCATION. As predicted, there was no statistically significant difference between the difficulty of detecting tempo modulation on beat 1 of the bar as opposed to tempo modulation beginning on beat 2 of the bar. However, contrary to predictions, this lack of difference also prevailed for rhythm modulation.

This result suggests that contrary to the argument put forward in Chapter 4, subjects may not have divided their attention between oscillators entrained to the beat-level time interval and oscillators entrained to the bar-level time interval as had been predicted. The failure to reject the null hypothesis in this case does not falsify the basic architecture of the model proposed in Chapter 3, so much as challenge the way in which the listener was expected to attend to the different components of the model.

As an adjunct to the above discussion of Hypothesis 3, it should also be mentioned that an unpredicted yet statistically significant interaction between DIRECTION and LOCATION was found in Experiment 2. Here, essentially, modulation of either type (rhythm or tempo) was easier to detect on beat 2 when it was early than on beat 1 when it was early. Early change on beat 1 would mean that the first beat of each bar would be lengthened (beat 1 would fall earlier than it should no matter whether the stimulus was undergoing rhythm or tempo modulation). In contrast, early change on beat 2 would mean that the first beat of each bar would be shortened (beat 2 would fall earlier that it should for either rhythm or tempo modulation). It may be the case that listeners are accustomed to hearing beat 1 lengthened in musical performances (see Clarke 1985, Todd 1985). If they heard beat 1 being shortened this would sound unusual and might, therefore, be more noticeable. One fact that argues against this explanation, however, is that if the length of beat 1 were an important factor in modulation detection, one would also expect to see a statistically significant difference between *late-beat1* and *late-beat2*, with *late-beat1* being more obvious owing to the shortening of beat 1. However, for Experiment 2, no statistically significant difference was found between *late-beat1* and *late-beat2*. It should also be noted that an explanation relying on the expected lengths of beats, however plausible, is not reducible to the workings of either the OSCILLATOR BANK or the SAS, and therefore neither supports nor falsifies the model outlined in Chapter 3.

TYPE by LOCATION in previous literature. The effect of the metrical location of the start of rhythm or tempo modulation has only been sparsely and indirectly addressed in the literature. As noted above, a number of performance studies have found that those onsets that are most displaced by rubato tend to fall on metrically important points in the bar (Clarke 1985; Shaffer 1981; Todd 1985, for example). However, these studies do not speak specifically about the *perception* of beat shortening or lengthening and so cannot be compared directly with the current results. Recall also that Wang (1984) recorded a highly conditional interaction concerning the effect of metrical location on tempo modulation detection: at the risk of simplifying Wang's results, when tempo modulation began on beats 1, 2 or 3 of a 4/4 bar, tempo deceleration was detected sooner than acceleration, however this trend was reversed when modulation began on beat 4. Although Wang's (1984) result is not easy to assimilate with the statistically significant DIRECTION by LOCATION interaction found in Experiment 2, both findings suggest that there is much that is not yet well understood about the influence of metrical location of time-interval modulation detection. Clearly, more experimentation is required.

Hypothesis 4

- *Hypothesis 4. There will be a statistically significant effect owing to PIECE in the rhythm-modulation conditions. Owing to the lack of beat subdivision, rhythm modulation in the Author piece will be more difficult to detect than equivalent modulation in the three pieces with beat subdivision. It is unknown how PIECE will affect tempo modulation perception.*

Experiment 3 alone was designed to test Hypothesis 4. As predicted, on the main effect PIECE, for the *Author* piece (which contained no beat subdivision) modulation detection (both rhythm and tempo) was statistically significantly more difficult than modulation detection in the *Bach*, *Chopin* or *Hindemith* pieces. Experiment 3, in striving for ecological validity, sacrificed some of the control of variables that had been present in Experiments 1 and 2. For this reason, the differences between level of the variable PIECE might be attributed to other uncontrolled sources of variance rather than the presence or absence of beat subdivision. One such source of variance might be the familiarity of melodic or harmonic style. However, this alternative fails to account for the greatest difference in means between pieces, namely the difference between the mean beats-to-respond for the *Bach* piece and the *Author* piece, since both of these pieces were in similar harmonic and melodic styles.

The original reason given in Chapter 4, for predicting that beat subdivision would aid modulation detection, was that phase resetting of the oscillators in the OSCILLATOR BANK would be more obvious (higher probability of resetting being detected) for oscillators entrained to a beat subdivision, than for oscillators entrained to the beat period. However, if one accepts the above explanation of Hypothesis 1 regarding Experiment 3, one must also accept that the perception of rhythm modulation may have been mediated by the SAS as well as the OSCILLATOR BANK. Recall that for Hypothesis 1, it was suggested that, perhaps, the greater complexity of the stimuli in Experiment 3 resulted in chunks of only a beat long being stored in the BUFFERS of the SAS. This would mean that direct beat-to-beat comparisons could also be carried out in the SAS as well as in the OSCILLATOR BANK. However, this use of two processes for rhythm modulation detection, would be less likely to occur in the unsubdivided *Author* piece. The *Author* piece's relative lack of complexity (lack of subdivision) would mean that beat-to-beat comparison would not be carried out in the SAS. The suggestion that modulation in the Author piece was detected using contrasting processes for rhythm and tempo modulation (as seems to have been the case for all stimuli in Experiments 1 and 2), is born out by the trends of the TYPE by DIRECTION by PIECE interaction for identification data shown in Figure 5.11. The TYPE by DIRECTION interaction for the Author piece shows an opposite trend from that of the other three pieces. The TYPE by DIRECTION interaction for the Author piece is more like the TYPE by DIRECTION interactions observed in Experiments 1 and 2.

Again, as was the case for Hypotheses 1 through 3, no previous studies make use of exactly the same variables as are employed in Experiment 3. Recall that two previous studies by Duke (1989a) and Kuhn (1987) had examined the effect of beat subdivision on detecting tempo differences between paired examples. Their only conclusion, however, seems to have been that beat subdivision caused subjects to confuse tempo (beat rate) with rhythmic activity (speed of beat subdivision) on a conceptual level. The only other study explicitly to address the topic of beat subdivision is Wang (1983), who as noted in Chapter 2, confounds the presence or absence of beat subdivision with three other factors (starting tempo, rate of tempo modulation and piece). Sadly, these confounds make her results uninterpretable, and therefore unfit for comparison with the current results. Given this confusion in the literature, and the fact that Experiment 3 used extracts of real keyboard music, future research concerning beat subdivision might start by examining the perception of simpler, purely rhythmic stimuli (such as were used in Experiments 1 and 2 here). This more basic research could serve to fill in the gaps that presently exist in the understanding of the role of metric phenomena in rhythm and tempo perception.

Learning

In reporting the results in Chapter 5 and in discussing them so far in this chapter, little attention has been paid to the fact that subjects were generally able to improve at the experimental task as each experiment progressed. In all three experiments, VERSION was a statistically significant main effect for beats-to-respond data. For identification data, VERSION was statistically significant as a main effect in Experiment 2 and approaching significance in Experiment 3. In all experiments, regardless of the type of data used as a dependent variable, the replication version was on average easier than the original version. This result betokens learning taking place over trials. Moreover, in all three experiments, specific statistically significant interactions between VERSION and other variables were observed, indicating that greater amounts of learning took place under some treatment conditions than under others. Several of these interactions have been mentioned in Chapter 5 but were too specific to be interpreted in any theoretically meaningful way. However, one interaction involving the variable VERSION is of particular musical interest, and has yet to be discussed in detail.

In Experiment 1 there was a statistically significant interaction between INITIAL BEAT RATE and VERSION for beats-to-respond data (refer back to Figure 5.11). In this case, learning took place at all initial beat rates apart from at 45 bpm, where the replication was marginally more difficult than the original. The greatest difference between original and replication conditions was for 190 bpm, but this difference is in part due to the high beats-to-respond mean of *190-original*. The lowest absolute cell mean for the whole interaction was *120-replication*. These results suggest that the greatest amount of learning took place at moderate to fast tempi (120-190 bpm). It is interesting that no learning took place at the slowest tempo (45 bpm). Learning here is specific to the experimental task, and any generalizations from these results should be made with extreme caution. Nevertheless, if these results are applicable to more musical situations, they suggest that learning to discriminate between time intervals (learning "rhythmic sensibility") is more efficacious at medium-fast tempi than at very slow tempi. Generalizing even further, if it can reasonably be assumed that one's ability to differentiate between rhythmic nuances in perception determines one's ability to do the same in performance, then practising nuanced rhythms such as swing or rubato will be less effective at slow tempi than at medium-fast tempi.

Methodological Considerations

The majority of the methodological considerations were addressed in Chapter 4 when the design of the three experiments was discussed. However, in the light of the results reported in Chapter 5, a few other issues now need to be considered.

Firstly, all three experiments reported low reliability coefficients for beats-to-respond data and especially identification data. Several explanations might be offered for this seeming lack of association between how a given subject performed on the first time of hearing a particular treatment condition, and how he/she performed on the second time of hearing that condition. Firstly, any type of reaction-time data (such as beats-to-respond data) can be expected to contain a considerable amount of error variance (Welford 1980; Ratcliff 1993). Four noteworthy sources of this error variance for the present studies can be assumed to be the following: (1) the maximum motoric speed at which subjects can push the response button, i.e., the kinematic limits of the subject's finger, (2) the subjects' motivation to press the response button as quickly as possible, (3) subjects' level of alertness or, conversely, fatigue and (4) subjects' level of concentration, i.e. the extent to which they attend to the experimental task. Naturally all of these factors will vary from subject to subject, and in one subject over time (during the course of one experimental session, for instance). Beats-to-respond data, resulting as it does from a speeded task, is vulnerable to all of the above types of error.

A further possible distortion of the beats-to-respond data could lie in the varying degrees of conservatism among subjects. In each experiment subjects were instructed to press the response button *as soon as* they detected any anomalous IOI, and not to concern themselves with classifying the type of modulation at that point (see the instructional rubrics in Appendix D). However, in actual fact, there was no way of knowing if a subject did indeed provide an immediate response, or whether he/she "wasted" time before pressing the response button trying to decide what type of modulation was occurring. Of the 84 subjects who participated across the three experiments, four intimated in informal conversations afterwards that they had sometimes waited a bar for confirmation that a modulation had taken place. Of course, the possible use of this strategy does not invalidate the results; the extent to which subjects need to wait for confirmation of a modulation is itself a measure of how difficult that modulation is to detect (and perceptual difficulty was what the present studies were intended to measure). Nevertheless, it should be born in mind that, beyond the four who admitted to hesitating, many more subjects may have delayed their response by a bar or two. In short, despite the wording of the instructions, the three experiments may not be entirely exempt from the type of speed/accuracy trade-off where conservative subjects will not react until they are absolutely sure that the to-be-detected event has taken place. This speed-accuracy trade-off is widely recognized in the literature as an inevitable feature of forced-choice reaction-time paradigms (Lachman, Lachman and Butterfield 1979; Luce 1986; Pachella 1974; Welford 1980).

Despite the above acknowledged limitations of reaction-time data, the reliabilty coefficients for this dependent variable were still higher than the reliability coefficients for identification data. In fact, the greater reliability of beats-to-respond data compared to identification data is puzzling since identification data, not being the result of a speeded task, should not be so vulnerable to the aforementioned sources of error variance.

A further, partial explanation for the low r^2 values for both data types may simply be that, as mentioned in Chapter 5, the presence of interactions between the variable VERSION and other variables. Such interactions would result from subjects exhibiting greater learning (a greater drop in beats-to-respond or wrong answers) between the original and replication of one condition than between the original and replication of another condition. Figure 6.2a illustrates this idea with a hypothetical set of results from Experiment 1 for subject Joe Average/Josephine Mean: beats-to-respond scores for original trials are given along the x-axis and beats-to-respond scores for replications of the same condition along the y-axis. For the first condition *rhythm-early-45* the replication version has a higher beats-to-respond value than the original; for the second condition, *rhythm-early-75* the replication and original version are roughly equal; for the third condition, *rhythm-early-190*, the replication version has a lower value than the original version. These results suggest that this subject exhibited the greatest amount of learning with stimuli at 190 bpm, there was no learning at 75 bpm, whilst stimuli at 45 bpm elicited "negative" learning (performance deteriorated across trials). The existence of this initial-beat-rate-specific learning means that the correlation coefficient between original and replication is lower than if the amount of learning had been constant between original and replication versions for all conditions (as in Figure 6.2b).

The last issue concerning reliability may be dealt with briefly. All three experiments revealed low r^2 values between the two data types. To recapitulate, the amount of mutual variance accounted for by the

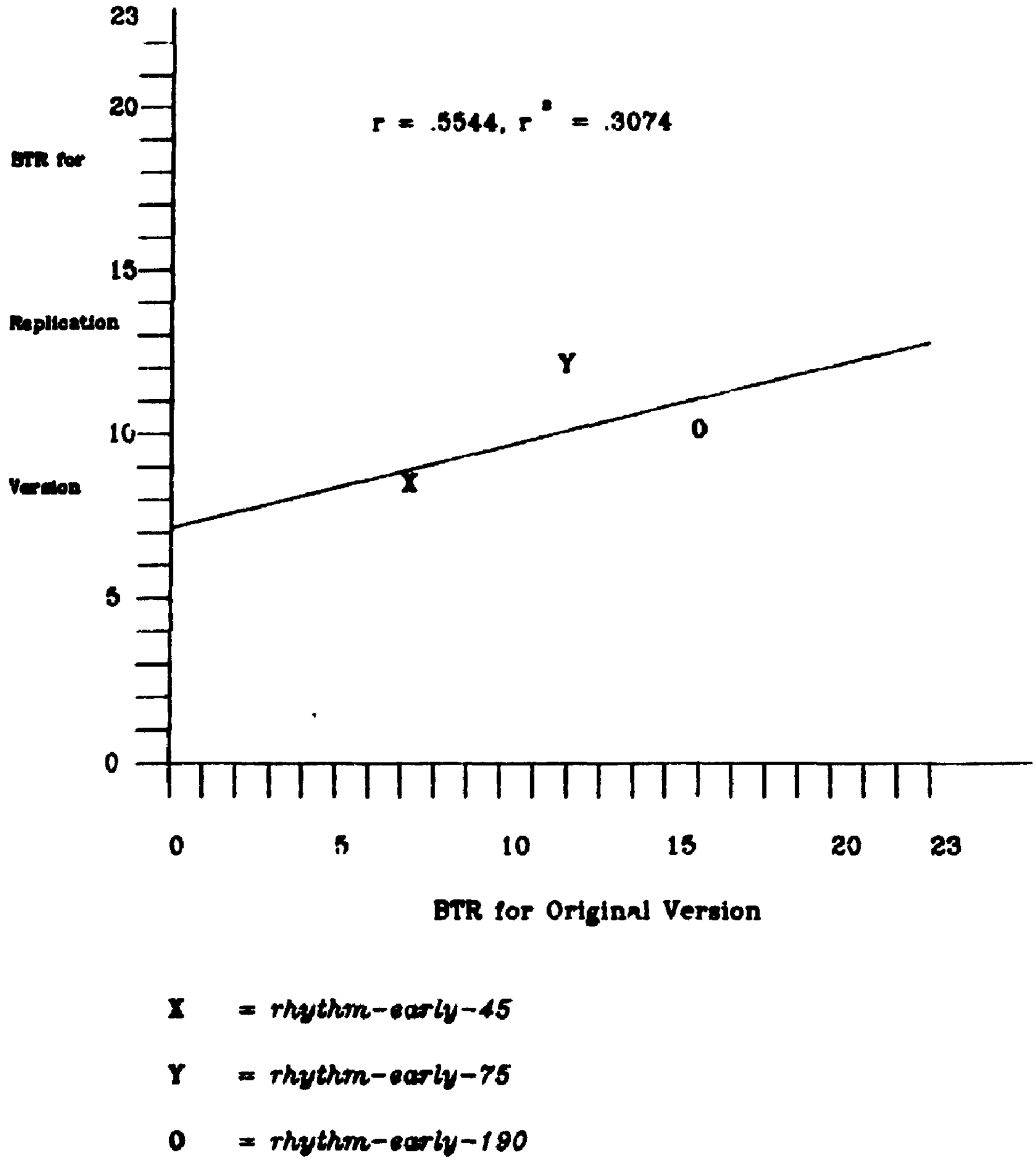


Figure 6.2a. Graph of original versus replication versions of the same conditions for a hypothetical subject. Here the reliability coefficient is low owing to differential learning across conditions.

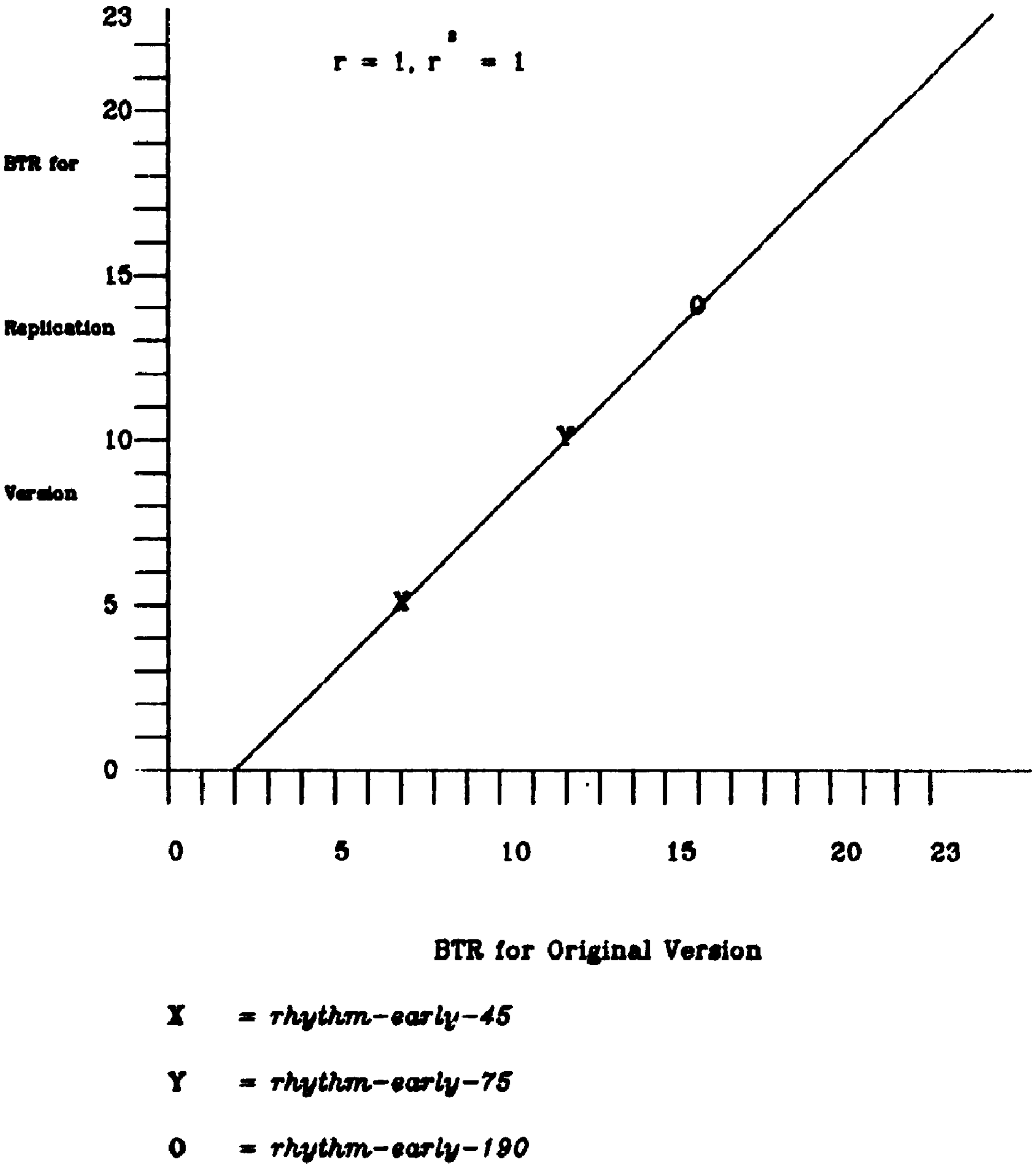


Figure 6.2b. Graph of original versus replication versions of the same conditions for a hypothetical subject. Here the reliability coefficient is very high owing to identical learning across conditions.

beats-to-respond and the identification data was as follows: Experiment 1 = 19%, Experiment 2 = 19%, Experiment 3 = 7%. Of these three, the result for Experiment 3 may be readily attributed to the overall lack of difference between conditions (lack of wrong answers) for identification data. As regards Experiments 1 and 2, the above meagre associations can be explained by highlighting the difference in trends across conditions for the two data types. The difference in trends has been touched on earlier in this chapter in the discussion of the four experimental hypotheses. Essentially, collection of the two data types may well entail the subject integrating stimulus information over contrasting time spans, making one task easier for certain treatment conditions than the other. Generally, it is an inevitability of rhythm-perception experimentation that, whilst one would ideally wish to measure the cognition and memory of time intervals between stimulus events instantaneously, the very means of measurement takes time and thus it is impossible to measure every stimulus time-interval on an exactly equal footing.

Conclusion

The Main Hypothesis of this dissertation, as it was originally formulated in Chapter 1 is as follows:

- Temporal changes that occur suddenly, i.e., within the span of the psychological present, are detected in perception; temporal changes that occur gradually, i.e., outside the span of the psychological present, are detected by comparing a memory trace with perception.

In general this hypothesis has been born out by the results of the three experiments. Rhythm modulation, as an example of sudden temporal change, exhibited no statistically significant difference in perceptual difficulty when the anomalous onset occurred early or late relative to the pattern established by previous events. In addition, rhythm modulation was easiest to detect at slow-medium tempi. In contrast, tempo modulation, as an example of gradual temporal change, exhibited a strong difference in perceptual difficulty, with a late anomalous onset (tempo deceleration) being easier to detect than a early anomalous onset (tempo acceleration). Also, in contrast to rhythm-modulation perception, tempo-modulation detection was most difficult at slow tempi and easiest at fast tempi (with the exception of 300 bpm for one data type). In Experiment 2 no differences were found between anomalous-IOI detection on beat 1 of the bar and the same on beat 2 of the bar for either rhythm or tempo modulation. In Experiment 3 the same trends were observed for both rhythm modulation and tempo modulation across the four excerpts of keyboard music used. Despite the lack of interaction between TYPE and LOCATION in Experiment 2, and the lack of interaction between TYPE and PIECE in Experiment 3, the highly significant interactions between TYPE and DIRECTION and TYPE and INITIAL BEAT RATE in Experiments 1 and 2 support the hypothesis that perception of rhythm modulation and tempo modulation are mediated by different mental processes. This conclusion, in turn, supports the yet more general conclusion that sudden temporal changes are detected via different mental processes from gradual temporal changes.

In Chapter 1, it was claimed that the problem of how temporal changes are perceived in various musical contexts had both a theoretical and a practical import. The theoretical reasons for tackling this problem were addressed in the model outlined in Chapter 3: to try to account for the multiplicity of experimental results reviewed in Chapter 2, to provide a model that paid heed to the general findings of time psychology, and to provide a model that worked event by event in "real" time. Positing the dual operation of a set of entrainable neural oscillators and a short auditory store allowed one, by and large, to realize these purposes. The practical import of the study lies in the way some of the results might be applied to everyday music listening. The results regarding rhythm modulation point to some of the ways local alterations of note lengths such as rubato, *inégalités* or swing might be perceived: the results here suggest that such rhythmic features would be most perceptible at medium tempi, that they are equally easy to perceive on beat 1 or beat 2 of the bar, and that beat subdivision facilitates their detection. The results regarding tempo modulation are suggestive of the manner in which less local temporal changes (from the bar level and longer) might be perceived: the results here suggest that gradual tempo modulation would be most easily perceived as tempo deceleration, when it begins at a fast tempo, and when the beat is subdivided. All of the aforementioned musical techniques are important features that add expression to a performance.

As has been hinted in previous paragraphs there remains much to be understood about the perception of rhythm and tempo modulation. Future research could fruitfully address itself to a number of different problems:

1. The possible interaction between time-interval modulation and melodic interval size. The literature on auditory stream segregation suggests that temporal relationships between notes (notably the perceived order of notes) are perceptually distorted when the notes sound rapidly and when the melodic interval between adjacent tones is wide (see Handel 1989 for a review of this literature). It would be of interest to know if large melodic intervals also distort listener's sense of relative time intervals. In other words, is detection of rhythm and tempo modulation made more difficult by larger melodic intervals and facilitated by narrow melodic intervals?

2. In all of the experiments in the present study, drum or piano sounds were used, sounds with quick or almost instantaneous onsets. However, some research suggests that with slower attack rates (longer than 100 ms) the perceptual onset of a tone is more ambiguous (Vos and Rasch 1981). For this reason, it would also be of interest to know if stimuli with slower note-attack rates make the detection of rhythmic anomalies more difficult than stimuli with percussive note-attacks. Such experimentation could use either synthesized tones, or for ecological validity, sampled instrument tones with slow attacks (e.g. bowed violin, or organ flute sounds).

3. The tantalizingly inconclusive results in Experiment 2 concerning the influence of metrical location on time-interval modulation detection beg further experimentation in this area. For instance, one might take an isochronous yet metrically unambiguous musical extract, and systematically lengthen and shorten different beats

of the bar to see which metrical location (if any) facilitates detection of these inter-onset anomalies.

4. It will be remembered that the model in Chapter 3 proposed that categorical rhythm perception is driven by the shortest time interval present in the stimulus at any given time. Therefore, if the shortest time interval is not integrally related to longer time intervals, the model predicts that categorical perception would be disrupted. It would be of value to test this prediction experimentally.

5. The existence of categorical perception might be tested in areas of time-interval perception other than just meter perception (Clarke 1987a). For example, it is entirely possible that there exist (for trained musicians at least) categories of tempo. In other words, the musical tempo range may not be a continuum, but may be divided up into discrete categories in the same way that the range of IOI ratios seems to be. If this were the case, discrimination between different tempi would be more acute when those different tempi straddle a categorical boundary, than when they lie within one category.

6. Finally, the model outlined in Chapter 3 might be tested, developed and refined by programming stages of it as a computer model or as a series of computer models. Owing to their mechanistic nature, the stages that seem at first blush to be most amenable to computer modelling are the OSCILLATOR BANK, and METER GATING.

It is hoped that this dissertation will provide a springboard for further enquiry into these and other fascinating problem areas to do with the perception of musical rhythm and tempo. However, the report provided in this and the foregoing chapters must suffice for now: answers to further questions lie in the province of the future.

APPENDIX A

PILOT STUDY

Before the three main experiments described in this dissertation were designed and carried out, a small-scale pilot study was run to investigate university music students' perception of tempo modulation. At the time, the purpose of this pilot was to ascertain to what extent tempo modulation of a given magnitude would be perceivable and whether there was likely to be any difference among initial-beat-rate conditions should a more extensive study be embarked upon. This Appendix describes the design, preparation, performance and results of the pilot study.

Design

In order to explore the types of variables that might be of interest in a more extensive study, this pilot was intended to examine the relative difficulty of perceiving tempo modulation in unpitched drum rhythms under certain conditions. Each drum rhythm was 24 beats long and three independent variables were manipulated:

1. DIRECTION. Tempo modulation occurred in one of three possible DIRECTIONS, the tempo got *faster*, it got *slower*, or there was no modulation resulting in a *constant* condition. When the rhythm got faster, it accelerated in a linear manner by a total of 20% from the initial beat rate over the span of fourteen beats. When the rhythm got

slower, it decelerated in a linear manner by a total of 20% from the initial beat rate over the span of fourteen beats. The constant (or no-modulation) condition was included as a foil and was discounted from statistical analysis. In trials that included a tempo modulation, the starting point of the modulation was randomly determined in order to discourage guessing or some other type of predictive strategy by the respondents. Figure A.1 gives examples of both accelerating tempo modulation and decelerating tempo modulation.

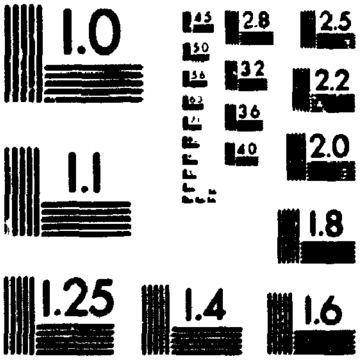
2. INITIAL BEAT RATE. INITIAL BEAT RATE had three levels, differentiated by the tempos of the isochronous drum-beats at the very beginning of each trial This INITIAL BEAT RATE is denoted by the value *t* in Figure A. Since, of course, beat rate is inversely proportional to time interval, the greater was *t* the shorter were the IOIs between the isochronous beats. Drum rhythms started at 45 bpm (initial IOI = 1333 ms), 105 bpm (initial IOI = 571 ms) or 245 bpm (initial IOI = 245 ms). These INITIAL BEAT RATES may be characterized as "very slow," "moderate" and "very fast" respectively. All three are shown in Figure A.1

3. METER. Dynamic accents were omitted from or added to selected notes within each drum rhythm resulting in three levels of the variable METER: *Unaccented* (no dynamic accents added), *Triple* (dynamic accents added every three beats), and *Random* (dynamic accents separated by an irregular and randomly determined number of beats).

The above independent variables resulted in a 3 (DIRECTION) by 3 (INITIAL BEAT RATE) by 3 (METER) design resulting in 27 trials for one experimental session. In addition, three practice items were played at

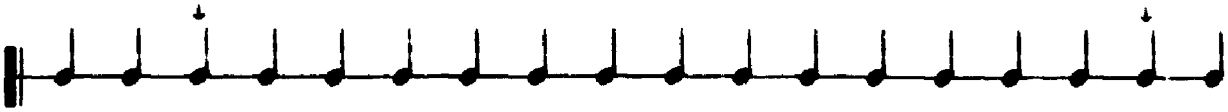
4 of/de 4

PM-1 3½"x4" PHOTOGRAPHIC MICROCOPY TARGET
NBS 1010a ANSI/ISO #2 EQUIVALENT



PRECISION℠ RESOLUTION TARGETS

a. t = 45 bpm.



45 45 44.4 (decelerating by 1.43% of t per beat) → 36 36

b. t = 45 bpm.

45 45 45.6 (accelerating by 1.43% of t per beat) → 54 54

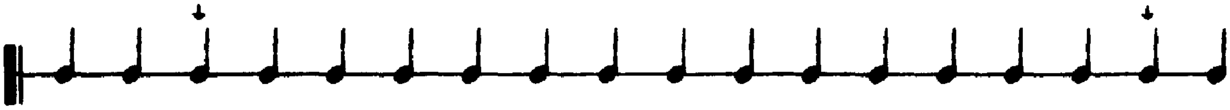
c. t = 105 bpm.

105 105 103.5 (decelerating by 1.43% of t per beat) → 84 84

d. t = 105 bpm.

105 105 106.5 (accelerating by 1.43% of t per beat) → 126 126

e. t = 245 bpm.



245 245 241.5 (decelerating by 1.43% of t per beat) → 196 196

f. t = 245 bpm.

245 245 248.5 (accelerating by 1.43% of t per beat) → 294 294

Figure A.1. Generic stimuli used in the pilot study. A.1a shows tempo deceleration at 45 bpm; A.1b shows tempo acceleration at 45 bpm; A.1c shows tempo deceleration at 105 bpm; A.1d shows tempo acceleration at 105 bpm; A.1e shows tempo deceleration at 245 bpm; A.1f shows tempo acceleration at 245 bpm. In each case t is the INITIAL BEAT RATE. The beginning and end of tempo change are marked by vertical arrows. The absolute location of the region of change was randomly determined for each trial. The change in tempo over time is indicated below each stave.

the beginning of each session to enable the subject to become familiarized with the experimental task.

Two dependent variables were garnered from each subject. On completion of each trial, subjects had two tasks:

1. To identify the direction of modulation, that is, "faster," "slower" or "constant." The identifications of the constant condition were ignored from statistical analysis, and the identifications of the two modulation conditions were treated as right/wrong answer data. Chi-square was used to analyze the this type of data, hereafter called "identification" data.

2. To rate the "obviousness" of the modulation on a scale of 0 to 4, 0 being "not at all obvious" and 4 being "very obvious." The three practice trials at the beginning of the experiment were designed to run the gamut of the "obviousness" scale. These practice trials, therefore, furnished subjects with some means of comparison when they came to the first proper trial. In the proper trials, if a subject misidentified the direction of modulation, then the obviousness rating for that particular trial was converted to a negative number. The reasoning behind this procedure was that if somebody gave a trial an obviousness rating of, say, 3 and misidentified the direction, then in fact they were more wrong in their judgment than somebody who also misidentified the same trial but who was unsure and only gave it an obviousness rating of 1. In addition, it would have been misleading to treat two obviousness ratings of 3 as equivalent data points if, in fact, one of those ratings entailed a correct identification and the

other an incorrect identification. ANOVA was used to analyze the obviousness data.

Preparation

All stimuli were generated on a computer and recorded to a high-fidelity chromium dioxide cassette tape for presentation to subjects. The rhythmic and dynamic aspects of the stimuli were generated using the Performer 3.6 Sequencer software from Mark of the Unicorn on an Apple MacIntosh Plus Computer. This software allowed definition of tempo gradations down to .1 of a metronome mark if desired, and 128 gradations of dynamic level. The sequencer drove a Yamaha DX7 II synthesizer set to an unpitched drum sound. The cassette machine used for recording was a Yamaha KX-800U Natural Sound Stereo Cassette Deck.

Subjects

Nine undergraduate and graduate students from the Music Faculty at the University of Western Ontario volunteered to participate as subjects in this experiment. Subjects were recruited by word-of-mouth and written advertisements. Music students were asked since they would be the population sampled in a more extensive study, and part of the reason for performing this pilot study was to ascertain how extensive tempo modulation should be in a stimulus neither to be too easy nor too difficult to detect.

Apparatus and Set-Up

Subjects were tested individually in a quiet, but not sound-proofed room. The experimental tape was played to subjects on a Yamaha KX-800U Natural Sound Stereo Cassette Deck through a pair of JVC HA-D330 Stereo Headphones. The output level of the tape deck was set to produce a comfortable and safe sound-pressure level.

The only other apparatus with which the subject had to interact was a pen and an answer sheet. Subjects were asked to make two marks on the answer sheet for each trial. Firstly they were asked to circle one of three letters to identify the direction of modulation "F" for "Faster," "S" for "Slower" or "C" for "Constant." Secondly, subjects were asked to circle a number from 0 to 4 to give the trial an obviousness rating.

Procedure

On entering the room, the subject was given a sheet of written instructions, and asked to read the details pertaining to the experimental task. While reading, the subject had the opportunity to ask the experimenter questions to clarify any points of procedure. Once the subject indicated comprehension of the instructions, the three practice items were played. Once the practice items were over, the tape was stopped, the experimenter verified that the subject understood the task and that the subject was happy with the sound pressure level. If all was in order the 27 trials proper were played in a random order.

To end the experimental session, the subject was thanked, paid $5 (Canadian), and debriefed. One session of the pilot study, including time for reading instructions and debriefing, was approximately 25 minutes.

Results

The results of the chi-square analysis that was performed on the identification data are shown in Table XXXIV. Tables XXXV through XXXVII show the accompanying cell contingencies. In these tables, and those that follow, the independent variables are represented by the following abbreviations: DIR = DIRECTION, IBR = INITIAL BEAT RATE and MET = METER.

TABLE XXXIV

PILOT STUDY, IDENTIFICATION DATA, CHI-SQUARE ANALYSIS

Independent Variable	Df	Chi-Square	p
DIR	1	14.063	.0002
IBR	2	11.625	.0030
MET	2	2.625	.2691

TABLE XXXV

PILOT STUDY, IDENTIFICATION DATA,
CONTINGENCY TABLE, MAIN EFFECT

DIRECTION

DIRECTION	Wrong	Right
Slower	1	80
Faster	17	64

TABLE XXXVI

PILOT STUDY, IDENTIFICATION DATA, CONTINGENCY TABLE, MAIN EFFECT

INITIAL BEAT RATE

INITIAL BEAT RATE	Wrong	Right
45	12	42
105	5	49
245	1	53

TABLE XXXVII

PILOT STUDY, IDENTIFICATION DATA, CONTINGENCY TABLE, MAIN EFFECT

METER

METER	Wrong	Right
Unaccent.	7	47
Triple	8	46
Random	3	51

Both DIRECTION and INITIAL BEAT RATE were statistically significant. Table XXXV shows that for DIRECTION the subjects detected tempo deceleration more easily than they did acceleration. Results for the factor INITIAL BEAT RATE show that subjects found tempo modulation detection easier the faster was the INITIAL BEAT RATE. As can be seen, there was no effect owing to METER.

The results of the Analysis of Variance carried out on the obviousness data are given in Table XXXVIII. The accompanying cells means are given in Tables IXL through XLII with the number of cases for each cell given in parentheses. Any cell means joined by a square bracket were not significantly different from each other at the .01 alpha level according to Duncan's multiple range test.

TABLE XXXVIII

PILOT STUDY, OBVIOUSNESS DATA, ANALYSIS OF VARIANCE

Source of Variance	Sum of Squares	Df	Mean Square	F	p
DIR	26.48	1	26.48	26.38	.001
Error	8.03	8	1.00		
IBR	21.06	2	10.53	14.31	.000
Error	11.77	16	.74		
MET	.30	2	.15	.33	.725
Error	7.28	16	.46		
DIR:IBR	14.93	2	7.46	12.34	.001
Error	9.68	16	.61		
DIR:MET	.74	2	.37	.88	.435
Error	6.78	16	.42		
IBR:MET	3.08	4	.77	2.57	.057
Error	9.59	32	.30		
DIR:IBR:MET	1.91	4	.48	1.74	.166
Error	8.81	32	.28		
Explained	68.51	17	4.03	8.46	.000
Residual	68.61	144	.48		
Total	137.12	161	.85		

TABLE IXL

PILOT STUDY, OBVIOUSNESS DATA, CELL MEANS OF MAIN EFFECT, DIRECTION

DIRECTION Slower	 Faster
1.65 (81)	0.84 (81)

TABLE XL

PILOT STUDY, OBVIOUSNESS DATA, CELL MEANS OF MAIN EFFECT, INITIAL BEAT RATE

	INITIAL BEAT RATE	
45	105	245
0.85 (54)	1.16 (54)	1.72 (54)

TABLE XLI

PILOT STUDY, OBVIOUSNESS DATA, CELL MEANS OF MAIN EFFECT, METER

	METER	
Unaccented	Triple	Random
1.29 (54)	1.26 (54)	1.19 (54)

TABLE XLII

PILOT STUDY, OBVIOUSNESS DATA, CELL MEANS OF 2-WAY INTERACTION, DIRECTION BY INITIAL BEAT RATE

DIR-ECTION	INITIAL BEAT RATE 45	 105	 245
Slower	1.69 (27)	1.33 (27)	1.93 (27)
Faster	0.02 (27)	0.98 (27)	1.52 (27)

As can be seen from tables XXXVIII and XXXVII, the obviousness data tell essentially the same story as the chi-square analysis. DIRECTION of tempo modulation was statistically significant with tempo decelerations being rated as more obvious than accelerations. INITIAL BEAT RATE was also statistically significant with faster INITIAL BEAT RATES given higher obviousness ratings. Thirdly, METER was not significant.

The ANOVA also reveals a statistically significant two-way interaction between DIRECTION and INITIAL BEAT RATE, as seen in Tables XXXVIII and XLII. For this interaction, while results show an ever-increasing obviousness rating the faster the INITIAL BEAT RATE for tempo *accelerations*, for tempo *decelerations* the same trend does not apply. Instead, for decelerations, the only statistically significant difference is between 105 and 245 bpm, with 245 bpm being deemed more obvious, whilst 45 bpm is marginally more obvious than 105 bpm and marginally less obvious than 245 bpm. This interaction is hard to interpret in any meaningful fashion.

Lastly in this results section, in order to measure to what extent high obviousness ratings were associated with correct identifications of modulation direction, a correlation coefficient was calculated between the two dependent variables. For this procedure, obviousness ratings were not converted to negative numbers, even if the corresponding identification was incorrect. The correlation coefficient was positive but not especially high, $r = .332$. From this $r^2 = .110$, meaning that one dependent variable only accounted for roughly 11% of the other's variance.

Discussion

Several noteworthy conclusions could be drawn from the above results. Firstly, the DIRECTION of tempo modulation affected the ease with which subjects were able to detect tempo modulation, with tempo deceleration being easier than tempo acceleration. These results support findings by Kuhn (1974) Madsen (1979) and Wang (1984). Since tempo deceleration involved a gradual lengthening of successive IOIs, it would be of interest in a more extensive study to see if IOI lengthening is detected more easily than IOI shortening in other rhythmic contexts besides tempo modulation.

Secondly, the INITIAL BEAT RATE of a stimulus before tempo modulation commenced was statistically significant, with faster stimuli facilitating tempo modulation detection. In a more extensive study it would be of interest to see if this trend would be replicated in a more detailed sampling of the musical tempo range. However, the aforementioned trend was tempered by the interaction between DIRECTION and INITIAL BEAT RATE. For whilst the trend held for tempo acceleration, it did not hold for tempo deceleration. For deceleration, the observation that modulation at 45 bpm was deemed marginally more obvious than modulation at 105 bpm (although the difference was not statistically significant) supported Ellis' (1991) findings that decelerations were easier to detect at slow tempi and accelerations were easier at fast tempi. However, the current results did not fully support Ellis' results since tempo deceleration at 245 bpm was deemed statistically significantly more obvious than deceleration at 105 bpm. Clearly more experimental data are required.

METER did not seem to affect tempo modulation detection. This finding was in line with Kuhn's (1987) study that found no significant difference in tempo modulation detection between melodies in triple and quadruple meter. For this reason, manipulation of meter as an independent variable (at least meter induced by dynamic accents) did not seem to provide a fruitful avenue for further empirical research.

Generally, judging from the low number of wrong answers--over all, only 18/162 (or 11%) of all modulations were incorrectly identified--subjects found the experimental task too easy. Therefore, to avoid the possibility of a ceiling effect, it was decided that in future the modulations should be more subtle, that is, with a lower overall percentage modulation over a greater number of beats.

In addition, in any future study, it seemed desirable to collect some type of data that measured subjects' perception of modulation as the modulation was occurring. Both the dependent variables in the current pilot study were limited to measuring the subject's short-term memory of events just after the stimulus had finished. All of the above points were later born in mind when it came to the design of the three experiments that make up the main part of this dissertation.

APPENDIX B

GEOMETRICAL PROOF OF EQUATION 3.2

In Chapter 3 the following equation was given to describe the relationship between the relative asynchrony of an IEIT pulse and the instantaneous energy level of an oscillator at the time-point of resetting:

$$AS_{rel} = .5(1 - EL_i/EL_{max}) \qquad (3.2)$$

where AS_{rel} is the relative asynchrony of an IEIT pulse compared to the already-established period of an entrained oscillator, EL_i is the oscillator's instantaneous energy level immediately before the IEIT pulse occurs, and EL_{max} is the oscillators maximum energy level (the energy level reached at the completion of one period of oscillation).

To prove equation 3.2, consider Figure B.1. The two right-angled triangles represent one complete cycle for a neural oscillator. Note that the two triangles are exact mirror images of each other since the ramp waveform of an oscillator is assumed to be perfectly symmetrical. The height of the line AZ represents the maximum energy level (EL_{max}) achieved at the completion of one cycle. The height of the line BY represents the instantaneous energy level of the waveform at the time-point of an asynchronous IEIT pulse (EL_i). The line segment AB represents the absolute asynchrony of the IEIT pulse compared to the cycle of the oscillator, therefore the ratio AB/2AC represents the

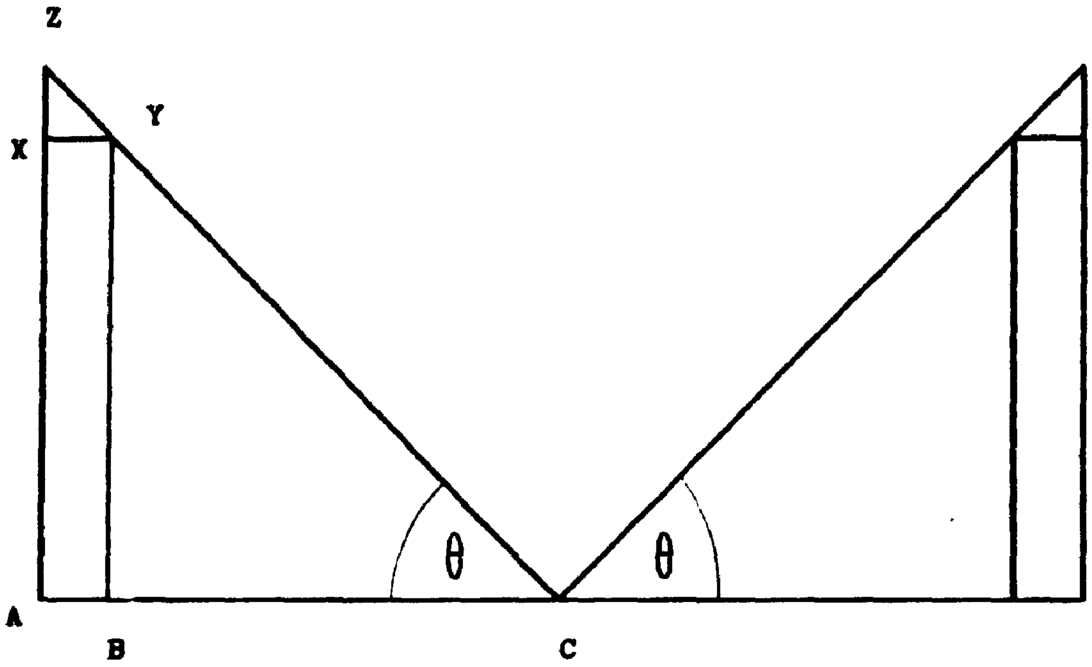


Figure B.1. Right-angled triangles representing one cycle of oscillation in a presumed neural oscillator. Note that the absolute length of the lines is not important for the purposes of proving equation 3.2.

relative asynchrony of the IEIT pulse expressed as a fraction of the already-established period of the oscillator (i.e., $AB/2AC = AS_{rel}$).

From Figure B.1 the following equations may be readily verified:

$$BY = AX$$

$$AX + XZ = AZ,$$

therefore, $$BY + XZ = AZ$$

therefore, $$BY/AZ + XZ/AZ = 1$$

therefore, $$\underline{XZ/AZ = 1 - BY/AZ} \qquad \text{(B.1).}$$

Also, $$\tan \theta = AZ/AC = BY/BC = XZ/XY$$

$$XY = AB$$

therefore, $$AZ/AC = XZ/AB$$

therefore, $$\underline{AB/AC = XZ/AZ} \qquad \text{(B.2).}$$

Combining equations B.1 and B.2,

$$AB/AC = 1 - BY/AZ$$

therefore, $$\underline{AB/2AC = .5\ (1 - BY/AZ)} \qquad \text{(B.3).}$$

Substituting the terms from the model outlined in Chapter 3, one gets,

$$AS_{rel} = .5\ (1 - EL_i/EL_{max})$$

which is equation 3.2.

APPENDIX C

RANDOM ORDERS OF TRIALS IN EXPERIMENTS 1, 2 AND 3

The following pages show the randomized orders of conditions as used in Experiments 1, 2 and 3. In each case, note that Order B is Order A in reverse. Only two stipulations were placed on the generation of these randomized orderings:

(1) Constant (no-change) conditions must be separated by at least one intervening trial from another constant condition.

(2) In Experiment 1, no two trials of the same INITIAL BEAT RATE were placed adjacent to each other; similarly, in Experiment 3, no two conditions of the same PIECE were placed adjacent to each other.

The "Start" column represents the randomly generated placements of the FAO (First Anomalous Onset), in either bar 2 or bar 3. The placement of the FAO was randomly varied to discourage guessing by the subjects.

As far as the variable VERSION is concerned, *original* is simply the first occurrence of a particular condition in a particular order, and *replication* is the second occurrence of the same condition in that order.

Experiment 1 - Randomized Orders

Trial # (Order A)	Condition	Start (Bar number)	Trial # (Order B)
1.	rhythm-early-190	2	45.
2.	rhythm-late-75	3	44.
3.	tempo-late-120	2	43.
4.	constant-45	-	42.
5.	rhythm-early-75	3	41.
6.	tempo-late-300	3	40.
7.	tempo-early-45	2	39.
8.	tempo-late-120	3	38.
9.	tempo-late-300	2	37.
10.	rhythm-late-45	3	36.
11.	rhythm-late-300	2	35.
12.	rhythm-late-190	3	34.
13.	tempo-early-75	3	33.
14.	rhythm-late-120	3	32.
15.	rhythm-late-45	2	31.
16.	tempo-early-120	2	30.
17.	constant-300	-	29.
18.	tempo-late-45	3	28.
19.	rhythm-early-300	2	27.
20.	tempo-early-75	2	26.
21.	tempo-late-190	2	25.
22.	rhythm-early-75	2	24.

Trial # (Order A)	Condition	Start (Bar number)	Trial # (Order B)
23.	rhythm-late-120	2	23.
24.	rhythm-early-300	3	22.
25.	rhythm-early-120	2	21.
26.	constant-75	-	20.
27.	tempo-early-300	3	19.
28.	constant-120	-	18.
29.	tempo-late-75	3	17.
30.	tempo-early-190	3	16.
31.	tempo-early-45	3	15.
32.	tempo-late-190	3	14.
33.	tempo-late-75	2	13.
34.	rhythm-late-300	3	12.
35.	tempo-early-120	3	11.
36.	rhythm-late-190	2	10.
37.	rhythm-early-45	2	9.
38.	tempo-early-190	2	8.
39.	tempo-early-300	2	7.
40.	tempo-late-45	2	6.
41.	rhythm-early-120	3	5.
42.	constant-190	-	4.
43.	rhythm-early-45	3	3.
44.	rhythm-early-190	3	2.
45.	rhythm-late-75	2	1.

Experiment 2 - Randomized Orders

Trial # (Order A)	Condition	Start (Bar number)	Trial # (Order B)
1.	rhythm-late-60-beat1	3	36.
2.	tempo-early-150-beat1	2	35.
3.	rhythm-late-150-beat2	2	34.
4.	rhythm-early-60-beat1	3	33.
5.	tempo-early-150-beat2	3	32.
6.	constant-60	-	31.
7.	tempo-late-150-beat2	2	30.
8.	tempo-late-60-beat1	3	29.
9.	rhythm-early-150-beat1	2	28.
10.	tempo-early-60-beat1	3	27.
11.	tempo-late-150-beat1	3	26.
12.	constant-150	-	25.
13.	rhythm-late-60-beat2	2	24.
14.	tempo-early-60-beat2	2	23.
15.	rhythm-late-150-beat2	3	22.
16.	tempo-late-60-beat2	3	21.
17.	rhythm-early-150-beat2	3	20.
18.	rhythm-late-150-beat1	2	19.
19.	rhythm-early-60-beat2	2	18.
20.	tempo-early-150-beat2	2	17.
21.	constant-60	-	16.
22.	rhythm-late-60-beat2	3	15.

Trial # (Order A)	Condition	Start (Bar number)	Trial # (Order B)
23.	rhythm-early-150-beat1	3	14.
24.	tempo-late-60-beat2	2	13.
25.	rhythm-late-150-beat1	3	12.
26.	tempo-early-60-beat2	3	11.
27.	tempo-late-150-beat1	2	10.
28.	rhythm-early-60-beat2	3	9.
29.	rhythm-late-60-beat1	2	8.
30.	constant-150	-	7.
31.	tempo-late-150-beat2	3	6.
32.	tempo-early-60-beat1	2	5.
33.	rhythm-early-150-beat2	2	4.
34.	tempo-late-60-beat1	2	3.
35.	tempo-early-150-beat1	3	2.
36.	rhythm-early-60-beat1	2	1.

Experiment 3 - Randomized Orders

Trial # (Order A)	Condition	Start (Bar number)	Trial # (Order B)
1.	rhythm-late-Bach	3	36.
2.	tempo-early-Hindemith	2	35.
3.	tempo-late-Chopin	2	34.
4.	constant-Hindemith	-	33.
5.	tempo-late-Author	3	32.
6.	rhythm-early-Hindemith	3	31.
7.	tempo-early-Chopin	3	30.
8.	tempo-early-Author	2	29.
9.	rhythm-late-Chopin	3	28.
10.	rhythm-early-Bach	2	27.
11.	rhythm-late-Author	2	26.
12.	tempo-early-Bach	2	25.
13.	rhythm-early-Author	3	24.
14.	tempo-late-Bach	3	23.
15.	tempo-late-Hindemith	2	22.
16.	constant-Bach	-	21.
17.	rhythm-late-Hindemith	3	20.
18.	rhythm-early-Chopin	3	19.
19.	rhythm-early-Author	2	18.
20.	tempo-late-Chopin	3	17.
21.	rhythm-late-Bach	2	16.
22.	tempo-early-Chopin	2	15.

Trial # (Order A)	Condition	Start (Bar number)	Trial # (Order B)
23.	rhythm-early-Hindemith	2	14.
24.	rhythm-early-Bach	3	13.
25.	rhythm-late-Hindemith	2	12.
26.	constant-Author	-	11.
27.	tempo-late-Hindemith	3	10.
28.	tempo-late-Bach	2	9.
29.	rhythm-late-Author	3	8.
30.	rhythm-early-Chopin	2	7.
31.	tempo-early-Hindemith	3	6.
32.	tempo-early-Bach	3	5.
33.	tempo-late-Author	2	4.
34.	constant-Chopin	-	3.
35.	tempo-early-Author	3	2.
36.	rhythm-late-Chopin	2	1.

APPENDIX D

INSTRUCTION AND RESPONSE SHEETS USED IN EXPERIMENTS 1, 2 AND 3

The ensuing pages reproduce the instruction and response sheets as they were printed for use in Experiments 1, 2 and 3. In each case, the only additions to the lay-out the subject saw are the page number and a lower-case Roman numeral for the purpose of identification with the following key:

i. Instruction rubric read by subject before taking part in Experiment 1.

ii. Instruction rubric read by subject before taking part in Experiment 2.

iii. Instruction rubric read by subject before taking part in Experiment 3.

iv. Response sheet used by subject in Experiment 1 to identify the type of change that had taken place in each trial.

v. Response sheet used by subject in Experiment 2 to identify the type of change that had taken place in each trial.

vi. Response sheet used by subject in Experiment 3 to identify the type of change that had taken place in each trial.

vii. Response sheet used by experimenter in Experiment 1 to note the reaction times in milliseconds for each trial.

viii. Response sheet used by experimenter in Experiment 2 to note the reaction times in milliseconds for each trial.

ix. Response sheet used by experimenter in Experiment 3 to note the reaction times in milliseconds for each trial.

i.

Dear Music Student,

thanks for participating in this study. This experiment is designed to see what types of changes in a steady drum rhythm are perceived more easily by music students.

The study is divided into three sections:

(1) You will hear five randomly spaced drum strikes. Your task is simply to press the response button as soon as you hear each strike.

(2) This is the main part of the study. You will hear two practice items and 45 trials consisting of drum rhythms. During each trial one of three things will happen:

(a) The tempo or speed of the rhythm will change (speed up or slow down). This is called a tempo change.

(b) The rhythm will stay the same tempo but it will get more and more uneven. This is called rhythm change.

(c) The rhythm will remain constant (neither tempo change nor rhythm change will occur).

In each trial you have two tasks:

Firstly, if a tempo or rhythm change occurs, push the response button as soon as you hear a change. At this stage do not worry about identifying the change type, simply push the button as soon as you hear that a note duration has been altered in some way. You will not need to push the button if the rhythm is constant.

Secondly, at the end of each trial, circle one of the three letters on the answer sheet to indicate what you think happened in the rhythm (Tempo change, Rhythm change, Constant).

Each trial will consist of random music (to clear your ears out), a voice announcing the trial number, the drum beats themselves, and 5 seconds silence for you to circle one of the letters on the answer sheet.

(3) Finally, you will hear five more drum strikes. Again, your task is to push the response button as soon as you hear each strike.

ii.

Dear Music Student,

thanks for participating in this study. This experiment is designed to see what types of changes in a steady drum rhythm are perceived more easily by music students.

The study is divided into three sections:

(1) You will hear five randomly spaced drum strikes. Your task is simply to press the response button as soon as you hear each strike.

(2) This is the main part of the study. You will hear two practice items and 36 trials consisting of drum rhythms. During each trial one of three things will happen:

(a) The tempo or speed of the rhythm will change (speed up or slow down). This is called a tempo change.

(b) The rhythm will stay the same tempo but it will get more and more uneven. This is called rhythm change.

(c) The rhythm will remain constant (neither tempo change nor rhythm change will occur).

In each trial you have two tasks:

Firstly, if a tempo or rhythm change occurs, push the response button as soon as you hear a change. At this stage do not worry about identifying the change type, simply push the button as soon as you hear that a note duration has been altered in some way. You will not need to push the button if the rhythm is constant.

Secondly, at the end of each trial, circle one of the three letters on the answer sheet to indicate what you think happened in the rhythm (Tempo change, Rhythm change, Constant).

Each trial will consist of random music (to clear your ears out), a voice announcing the trial number, the drum beats themselves, and 5 seconds silence for you to circle one of the letters on the answer sheet.

(3) Finally, you will hear five more drum strikes. Again, your task is to push the response button as soon as you hear each strike.

iii.

Dear Music Student,

thanks for participating in this study. This experiment is designed to see what types of tempo and rhythm changes in selected piano pieces are perceived more easily by music students.

The study is divided into three sections:

(1) You will hear five randomly spaced piano tones. Your task is simply to press the response button as soon as you hear each tone.

(2) This is the main part of the study. You will hear two practice items and 36 trials consisting of short excerpts of piano music. During each trial one of three things will happen:

(a) The tempo or speed of the music will change (speed up or slow down). This is called a <u>tempo change</u>.

(b) The rhythm will stay the same tempo but it will get more and more uneven. This is called <u>rhythm change</u>.

(c) The rhythm will remain <u>constant</u> (neither tempo change nor rhythm change will occur).

In each trial you have two tasks:

Firstly, if a tempo or rhythm change occurs, push the response button as soon as you hear a change. At this stage do not worry about identifying the change type, simply push the button as soon as you hear that a note duration has been altered in some way. You will not need to push the button if the rhythm is constant.

Secondly, at the end of each trial, circle one of the three letters on the answer sheet to indicate what you think happened in the rhythm (<u>T</u>empo change, <u>R</u>hythm change, <u>C</u>onstant).

Each trial will consist of random music (to clear your ears out), a voice announcing the trial number, the excerpt of piano music itself, and 5 seconds silence for you to circle one of the letters on the answer sheet.

(3) Finally, you will hear five more piano tones. Again, your task is to push the response button as soon as you hear each tone.

iv.

Response Sheet

Trial format: Random Music -- Voice -- Drum Beats -- Silence

Trial #	Change Type Tempo	Rhythm	Constant
Practice	T	R	C
Practice	T	R	C
1.	T	R	C
2.	T	R	C
3.	T	R	C
4.	T	R	C
5.	T	R	C
6.	T	R	C
7.	T	R	C
8.	T	R	C
9.	T	R	C
10.	T	R	C
11.	T	R	C
12.	T	R	C
13.	T	R	C
14.	T	R	C
15.	T	R	C
16.	T	R	C
17.	T	R	C
18.	T	R	C
19.	T	R	C
20.	T	R	C

Trial #	Change Type		
	Tempo	Rhythm	Constant
21.	T	R	C
22.	T	R	C
23.	T	R	C
24.	T	R	C
25.	T	R	C
26.	T	R	C
27.	T	R	C
28.	T	R	C
29.	T	R	C
30.	T	R	C
31.	T	R	C
32.	T	R	C
33.	T	R	C
34.	T	R	C
35.	T	R	C
36.	T	R	C
37.	T	R	C
38.	T	R	C
39.	T	R	C
40.	T	R	C
41.	T	R	C
42.	T	R	C
43.	T	R	C
44.	T	R	C
45.	T	R	C

V.

Response Sheet

Trial format: Random Music -- Voice -- Drum Beats -- Silence

Trial #	Change Type		
	Tempo	Rhythm	Constant
Practice	T	R	C
Practice	T	R	C
1.	T	R	C
2.	T	R	C
3.	T	R	C
4.	T	R	C
5.	T	R	C
6.	T	R	C
7.	T	R	C
8.	T	R	C
9.	T	R	C
10.	T	R	C
11.	T	R	C
12.	T	R	C
13.	T	R	C
14.	T	R	C
15.	T	R	C
16.	T	R	C
17.	T	R	C
18.	T	R	C
19.	T	R	C
20.	T	R	C

Trial #	Change Type Tempo	Rhythm	Constant
21.	T	R	C
22.	T	R	C
23.	T	R	C
24.	T	R	C
25.	T	R	C
26.	T	R	C
27.	T	R	C
28.	T	R	C
29.	T	R	C
30.	T	R	C
31.	T	R	C
32.	T	R	C
33.	T	R	C
34.	T	R	C
35.	T	R	C
36.	T	R	C

vi.

<u>Response Shee</u>t+

Trial format: Random Music -- Voice -- <u>Piano Music</u> -- Silence

<u>Trial #</u>	<u>Change Type</u> <u>T</u>empo	<u>R</u>hythm	<u>C</u>onstant
Practice	T	R	C
Practice	T	R	C
Practice	T	R	C
1.	T	R	C
2.	T	R	C
3.	T	R	C
4.	T	R	C
5.	T	R	C
6.	T	R	C
7.	T	R	C
8.	T	R	C
9.	T	R	C
10.	T	R	C
11.	T	R	C
12.	T	R	C
13.	T	R	C
14.	T	R	C
15.	T	R	C
16.	T	R	C
17.	T	R	C
18.	T	R	C
19.	T	R	C
20.	T	R	C

Trial #	Change Type		
	Tempo	Rhythm	Constant
21.	T	R	C
22.	T	R	C
23.	T	R	C
24.	T	R	C
25.	T	R	C
26.	T	R	C
27.	T	R	C
28.	T	R	C
29.	T	R	C
30.	T	R	C
31.	T	R	C
32.	T	R	C
33.	T	R	C
34.	T	R	C
35.	T	R	C
36.	T	R	C

vii.

Reaction Times (in ms)

Identification:
Pre-test drum strikes: (1)__________ (2)__________

(3)__________ (4)__________ (5)__________

Main Trials

1. __________
2. __________
3. __________
4. __________
5. __________
6. __________
7. __________
8. __________
9. __________
10. __________
11. __________
12. __________
13. __________
14. __________
15. __________
16. __________
17. __________
18. __________
19. __________
20. __________
21. __________
22. __________
23. __________
24. __________
25. __________
26. __________
27. __________
28. __________
29. __________
30. __________
31. __________
32. __________
33. __________
34. __________
35. __________
36. __________
37. __________
38. __________
39. __________
40. __________
41. __________
42. __________
43. __________
44. __________
45. __________

Post-test drum strikes: (6)__________ (7)__________ (8)__________

(9)__________ (10)__________

viii. **Reaction Times (in ms)**

Identification:
Pre-test drum strikes: (1)__________ (2)__________

(3)__________ (4)__________ (5)__________

Main Trials

1. __________
2. __________
3. __________
4. __________
5. __________
6. __________
7. __________
8. __________
9. __________
10. __________
11. __________
12. __________
13. __________
14. __________
15. __________
16. __________
17. __________
18. __________
19. __________
20. __________
21. __________
22. __________
23. __________
24. __________
25. __________
26. __________
27. __________
28. __________
29. __________
30. __________
31. __________
32. __________
33. __________
34. __________
35. __________
36. __________

Post-test drum strikes: (6)__________ (7)__________ (8)__________

(9)__________ (10)__________

ix. **Reaction Times (in ms)**

Identification:
Pre-test piano tones: (1)__________ (2)__________

(3)__________ (4)__________ (5)__________

Main Trials

1. __________
2. __________
3. __________
4. __________
5. __________
6. __________
7. __________
8. __________
9. __________
10. __________
11. __________
12. __________
13. __________
14. __________
15. __________
16. __________
17. __________
18. __________
19. __________
20. __________
21. __________
22. __________
23. __________
24. __________
25. __________
26. __________
27. __________
28. __________
29. __________
30. __________
31. __________
32. __________
33. __________
34. __________
35. __________
36. __________

Post-test piano tones: (6)__________ (7)__________ (8)__________

(9)__________ (10)__________

APPENDIX E

DATA CONVERSION

The purpose of this Appendix is to explain and exemplify how raw reaction-time data in milliseconds were converted into beats-to-respond data.

Once three or four subjects had each sat through an experimental session, the raw reaction-time data in milliseconds were encoded on computer, forming a single variable column in an SPSS/pc data file (Norušis 1991). To convert these data points into beats-to-respond data, the data file was then processed by an SPSS command file of over 300 lines in length. To exemplify how the command file processed the reaction-time data, consider the durations inherent in the stimulus shown in Figure E.1, which shows part of the *rhythm-late-75* condition from Experiment 1. In the ensuing description, "DATAREAC" is the name given to the raw reaction-time datum in milliseconds (the input to the command file), and "MODREAC" is the name given to the corresponding beats-to-respond datum (the output of the command file). The names DATAREAC and MODREAC were those used in the command file.

In all experimental stimuli, the timer began counting at the first anomalous onset (FAO). For the stimulus shown in Figure E.1, if a reaction-time of under 776 milliseconds (the duration of the first beat after the FAO) was recorded, then this value represented a beats-to-

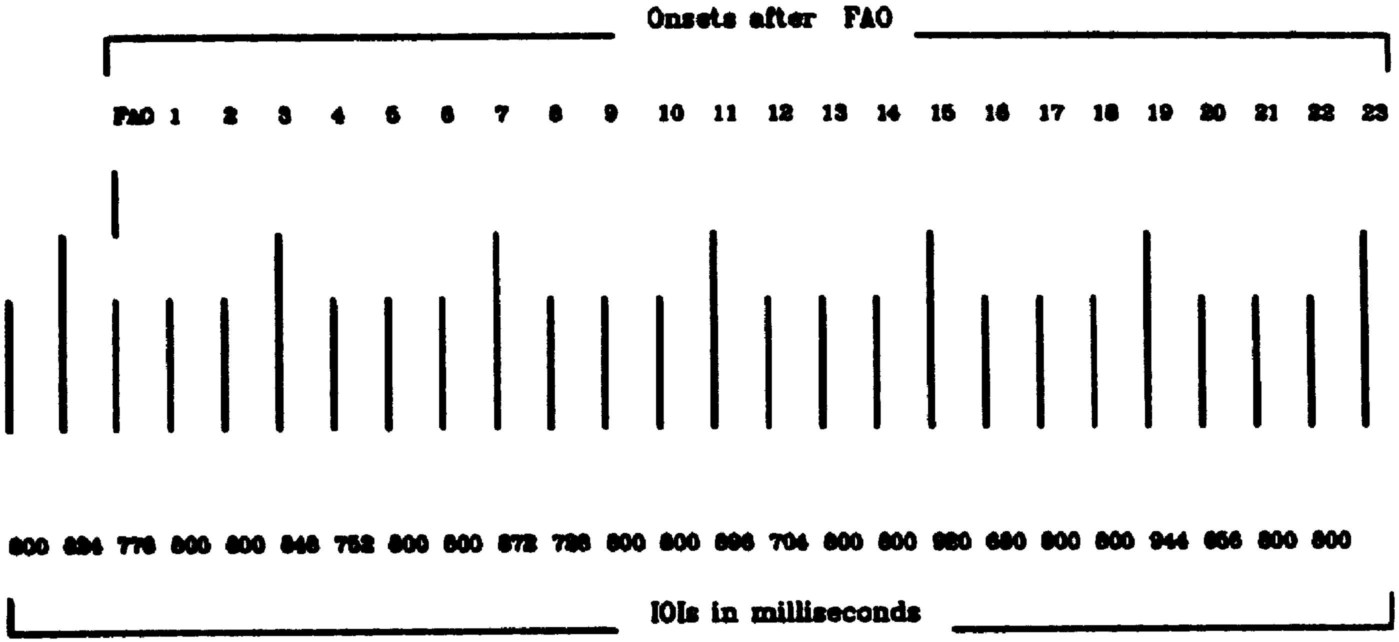


Figure E.1. IOIs in milliseconds for the rhythm-late-75condition. Beats are numbered from the FAO The FAO could be either in the second or third bar of the stimulus. Downbeats are indicated by longer vertical lines.

respond datum of only a fraction of one beat. Therefore the first expression in the conversion command file was:

IF (DATAREAC <= 776) MODREAC = DATAREAC/776.

In other words, if the reaction-time datum was less than or equal to 776 milliseconds, the beats-to-respond datum equalled the reaction-time datum divided by 776. For example a reaction time of 388 ms would be 388/776 or .5 beats.

If a reaction time of over 776 milliseconds was recorded, say 1226 ms, then its conversion to a beats-to-respond datum was more long-winded. A value of 1226 ms could be divided into two components: 776 milliseconds of it equalled 1 beat, whilst the remainder would be some fraction of the next beat, which in Figure E.1 is 800 ms long. Therefore 1226 ms in beats is

1 + (1226-776)/800 = 1.56 beats after the FAO.

Furthermore, because the beat durations at 1-2 and 2-3 are the same, that is, 800 ms, the same procedure could be applied to a reaction-time datum that lay between 1576 (776 + 800) and 2376 (776 + 800 + 800) milliseconds. For example, a reaction time of 2001 ms equalled 1 + (2001-776)/800 = 2.53 beats.

Therefore the second expression in the conversion command file was:

IF (DATAREAC > 776 AND <= 2376) MODREAC = 1 + (DATAREAC-776)/800.

In other words, if the reaction-time datum was between 776 and 2376 ms, its beats-to-respond value could be computed by subtracting 776 from the reaction-time datum, expressing the remainder as a fraction of the new beat duration (800 ms), and adding 1 (the number of beats already elapsed since the FAO) to the result.

One further example will suffice for the purposes of illustration. For any reaction-time datum greater than 2376 ms but less than 3224 ms (2376 + 848 [the duration at 3-4]), say 3005 ms, "3 and a bit" beats would have elapsed since the FAO. To compute the beats-to-respond datum exactly one would compute the following:

3 + (3005-2376)/848 = 3.74 beats.

Therefore, the third expression in the conversion command file was:

```
IF (DATAREAC > 2376 AND <= 3224) MODREAC = 3 + (DATATREAC-2376)/848.
```

In other words, if the reaction-time datum was between 2376 and 3324 ms, its beats-to-respond value could be computed by subtracting 2376 from the reaction-time datum, expressing the remainder as a fraction of the new beat duration (848 ms), and adding 3 (the number of beats already elapsed since the FAO) to the result.

As a general rule, then, the command file subtracted the number of milliseconds that had elapsed at the time-point of the last complete beat, expressed the remainder as a fraction of the current beat duration, and then added the integer value of the number of beats that had elapsed. It therefore becomes evident that a new expression was needed in the command file every time the beat duration changed. Moreover, every TYPE:DIRECTION:INITIAL-BEAT-RATE condition required a unique set of expressions for data conversion. Therefore, it can readily be seen how the overall command file for data conversion for each experiment was very long (particularly for Experiment 1 which had five INITIAL-BEAT-RATE conditions).

In the interests of brevity only the complete set of expressions for the condition *rhythm-late-75* from Experiment 1 is given as an example below. Note that the final expression converts any reaction-time datum greater than 18376 milliseconds (23 beats after the FAO) to a fixed value of 23 beats.

```
IF (DATAREAC <= 776) MODREAC = DATAREAC/776.
IF (DATAREAC > 776 AND <= 2376) MODREAC = 1 + (DATAREAC-776)/800.
IF (DATAREAC > 2376 AND <= 3224) MODREAC = 3 + (DATATREAC-2376)/848.
IF (DATAREAC > 3224 AND <= 3976) MODREAC = 4 + (DATAREAC-3224)/752.
IF (DATAREAC > 3976 AND <= 5576) MODREAC = 5 + (DATAREAC-3976)/800.
IF (DATAREAC > 5576 AND <= 6448) MODREAC = 7 + (DATAREAC-5576)/872.
IF (DATAREAC > 6448 AND <= 7176) MODREAC = 8 + (DATAREAC-6448)/728.
IF (DATAREAC > 7176 AND <= 8776) MODREAC = 9 + (DATAREAC-7176)/800.
IF (DATAREAC > 8776 AND <= 9672) MODREAC = 11 + (DATAREAC-8776)/896.
IF (DATAREAC > 9672 AND <= 10376) MODREAC = 12 + (DATAREAC-9672)/704.
IF (DATAREAC > 10376 AND <= 11976) MODREAC = 13 + (DATAREAC-10376)/800.
IF (DATAREAC > 11976 AND <= 12896) MODREAC = 15 + (DATAREAC-11976)/920.
IF (DATAREAC > 12896 AND <= 13576) MODREAC = 16 + (DATAREAC-12896)/680.
IF (DATAREAC > 13576 AND <= 15176) MODREAC = 17 + (DATAREAC-13576)/800.
IF (DATAREAC > 15176 AND <= 16120) MODREAC = 19 + (DATAREAC-15176)/944.
IF (DATAREAC > 16120 AND <= 16776) MODREAC = 20 + (DATAREAC-16120)/680.
IF (DATAREAC > 16776 AND <= 18376) MODREAC = 21 + (DATAREAC-16776)/800.
IF (DATAREAC > 18376) MODREAC = 23.
```

APPENDIX F

CHECK FOR TAPE SPEED CONSISTENCY

Given that timing and duration were of central importance to the experiments performed in this study (both as independent and dependent variables), one might seek assurance that the tape machine on which the stimuli were presented did not vary in speed to any appreciable degree. Although the tape machine used was of high quality (a Yamaha KX - 800U discrete head stereo cassette deck) as were the tapes (Sony CrO2 Cassette tapes), it seemed desirable to carry out further checks on the consistency of playing speed of the trials for all three experiments.

Procedure

To check tape speed, the Hewlett-Packard timer was connected to the tape machine in the manner outlined in Figure F.1. This setup differed from that used in the experimental sessions in that the left channel (containing the stimulus music) was connected to start the timer, and the right channel (containing the trigger signal) was connected to stop the timer. With the set-up in Figure F.1, if a typical trial was played on the tape machine, the following happened: if the clock was reset in the silence just before the stimulus began, the timer was started by the onset of the stimulus and stopped by the trigger signal (as illustrated in Figure F.2). This measurement could

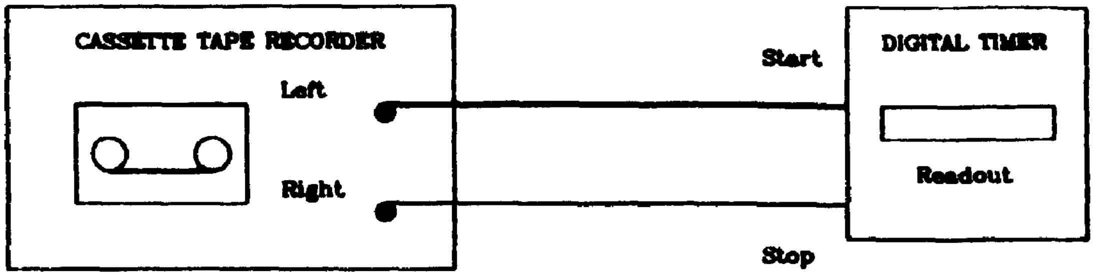


Figure F.1. Set-up of apparatus to measure tape speed. The left channel (stimulus onset) starts the timer; the right channel (trigger signal) stops the timer. In this way multiple samples of tape playing time can be taken from each tape.

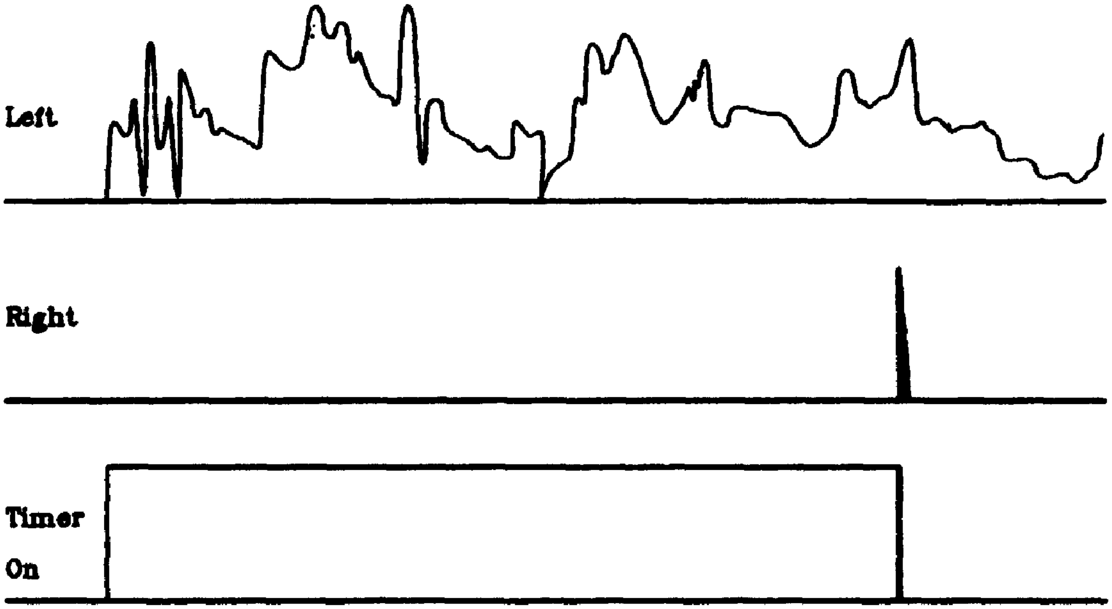


Figure F.2. Timer activity is response to the set-up shown in Figure F.1. The timer is started by the signal on the left channel and is stopped by the trigger on the right channel. The final readout on the timer gives the time the tape took to play a segment of known duration. The observed duration can then be compared with the theoretical duration to determine how fast or slow the tape is running be taken in turn for each trial on the tape resulting in a number of samples of playing time from stimulus onset to time of trigger signal taken from different positions on the tape. These measured times could be compared with "theoretical" onset-to-trigger times calculated from the durations generated by the sequencer program. To compare calculated and observed times, a ratio of observed/calculated could then be computed for each trial. An observed/calculated onset-to-trigger duration ratio of less than 1 would mean that the tape was running fast. Conversely an observed/calculated time ratio of more than 1 would mean that the tape was running slow.

The measurement procedure described above was carried out twice for both orders of stimuli for all three experiments. The first measurement for a given experiment (both orders) was performed the first day the experiment was run, immediately after the tape machine had been turned on. The second measurement for a given experiment (again both orders) was performed the last day the experiment was run, after all the experimental sessions had been completed for that day. In short, for the running of each experiment (which typically took 2 to 3 weeks) the very first and the very last task performed was the measurement of tape speed. In this manner four questions relating to potential tape-speed deviation could be answered: (1) by how much did the tape speed deviate from the "correct" calculated speed? (2) did the tape speed vary as a function of the machine "warming up" over one experimental session? (3) did the tape machine vary in speed across experimental orders? (4) did the tape speed vary as function of wear and tear over several weeks use?

Extent of Tape-Speed Variation

The variation in tape speed was monitored by computing the observed/calculated ratio for onset-to-trigger durations for each experiment separately.

In Experiment 1 the onset-to-trigger duration was measured twice for each of the 40 trials that contained a trigger signal, for orders A and B, giving 2 X 40 X 2 or 160 measurements in total. These measurements were each divided by the corresponding calculated onset-to-trigger value, giving 160 ratios. Descriptive statistics for the computed ratios for Experiment 1 are as follows:

Mean = 1.0014

Minimum = .994 Maximum = 1.008

Median = 1.0010 25%ile = .9997 75%ile = 1.0031

On average then, for Experiment 1 the tape machine was running 0.14% slower than the correct speed. Across trials and orders it varied by no more than +/-0.8% around the correct value (from minimum to maximum), and 50% of the onset-to-trigger durations (the interquartile range) were within +/-0.31% of the correct value. Taking the observed median as the central value, across trials and orders the speed varied by less than +/-0.7% (from minimum to maximum), and 50% of the onset-to-trigger durations (from the interquartile range) were within +/-0.2% of the median.

In Experiment 2 the onset-to trigger duration was measured twice for each of 32 trials that contained a trigger signal, for orders A and B, giving 2 X 32 X 2 or 128 measurements in total. These measurements were each divided by the corresponding calculated onset-to-trigger value, giving 128 ratios. Descriptive statistics for the computed ratios for Experiment 2 are as follows:

Mean = 1.0019

Minimum = .999 Maximum = 1.008

Median = 1.0017 25%ile = 1.0011 75%ile = 1.0027

On average, for Experiment 2 the tape machine was running 0.19% slower than the correct speed. Across trials and orders it varied by no more than 0.8% around the correct value (from minimum and maximum), and 50% of the onset-to-trigger durations (from the interquartile range) were within +0.3% of the correct value. Taking the observed median as the central value, across trials and orders the speed varied by less than +/-0.7% (from minimum to maximum), and 50% of the onset-to-trigger durations (from the interquartile range) were within +/-0.1% of the median.

In Experiment 3 the onset-to trigger duration was measured twice for each of 32 trials that contained a trigger signal, for orders A and B, giving 2 X 32 X 2 or 128 measurements in total. These measurements were each divided by the corresponding calculated onset-to-trigger value, giving 128 ratios. Descriptive statistics for the computed ratios for Experiment 3 are given as follows:

Mean = .995

Minimum = .991 Maximum = 1.001

Median = .9949 25%ile = .9938 75%ile = .9960

On average, in Experiment 3 the tape machine was running 0.5% faster than the correct speed. Across trials and orders it varied by no more than 0.9% around the correct value (from minimum and maximum), and 50%

of the onset-to-trigger durations (from the interquartile range) were within +/-0.7% of the correct value. Taking the observed median as the central value, across trials and orders the speed varied by less than +/-0.7% (from minimum to maximum), and 50% of the onset-to-trigger durations (from the interquartile range) were within +/-0.2% of the median.

The above statistics betoken an extremely consistent playing speed across all three experiments. In all cases deviations not only from the correct onset-to-trigger durations but also from the observed median are consistently less than +/-1%. In the vast majority of measurements the deviations are less than +/-0.5% around the correct value and the median.

Randomness of Tape-Speed Variation

Next, one might want to ascertain whether these small deviations around the median are randomly distributed across trials, orders and measurement sessions. If deviations are systematically ordered then one would expect long runs of consecutive observed/calculated ratios to be consistently above or below the median ratio. Figure F.3 illustrates three potentially damaging types of systematic tape-speed deviation. In Figure F.3a the tape speed changes consistently in one direction (decelerates in this case) over trials. As can be seen, the first 50% of the trials are below the median ratio, whereas the second 50% are above the median ratio. In Figure F.3b the tape speed is slower for order B than for order A. If ratios for orders A and B are chained together one would see a long run of ratios for order A below the median

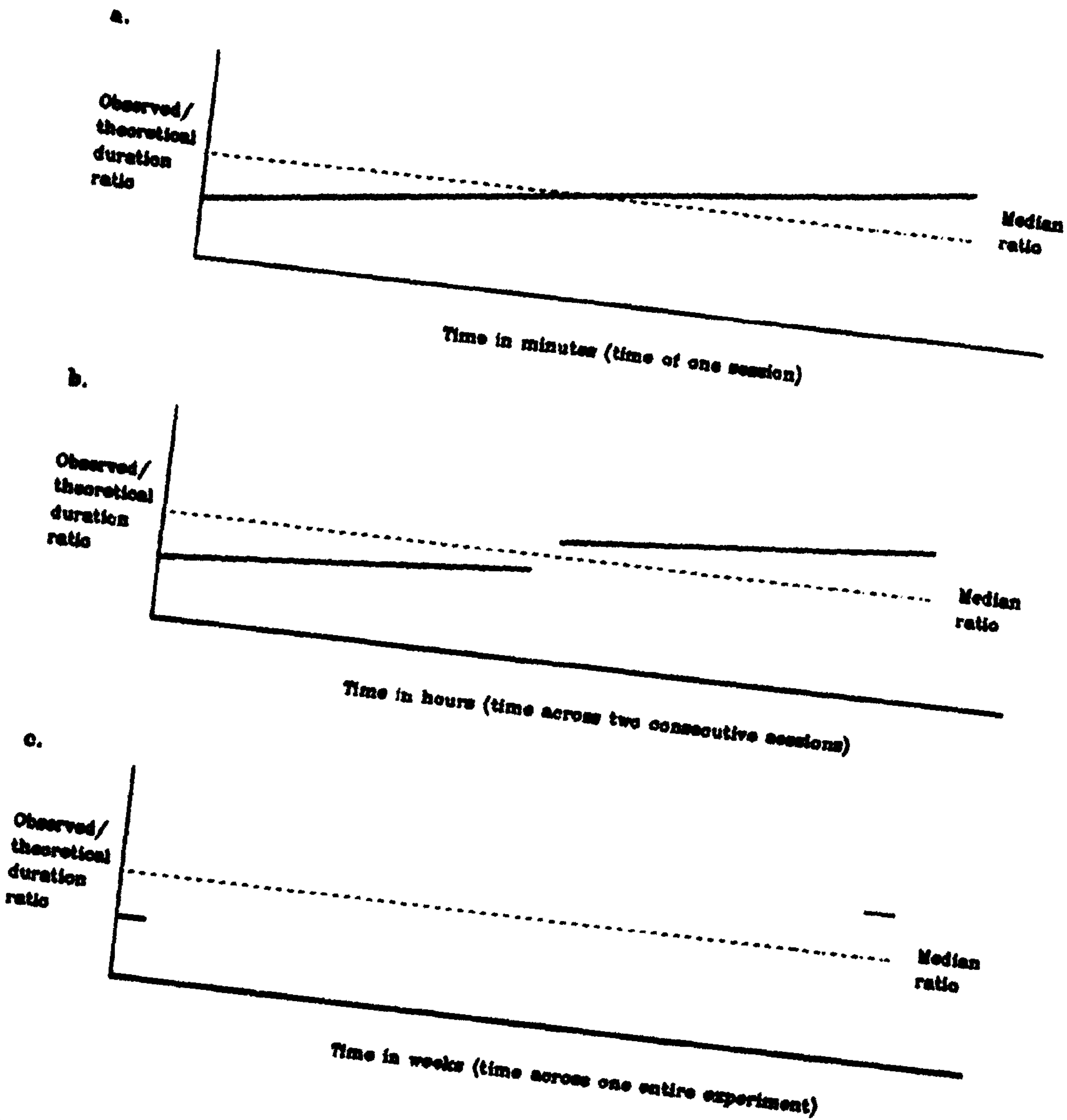


Figure F.3. Three scenarios with varying tape speed. Figure F.3a shows tape speed decelerating continuously over one experimental session. Figure F.3b shows tape speed running slower between two consecutive experimental sessions. Figure F.3c shows tape speed slowing over the period of an entire experiment ratio, and a long run of ratios for order B above the median ratio. In Figure F.3c the tape speed is slower for the measurement session performed on the last day of the experiment than it is on the first day. Again in this scenario one would see a long run of ratios below the median for the first day and a run of ratios above the median for the last day.

For each experiment observed/calculated duration ratios were arranged in the following order:

first day - order A - trials 1 to N
first day - order B - trials 1 to N
last day - order A - trials 1 to N
last day - order B - trials 1 to N.

To gauge the randomness of tape speed deviations around the median observed/calculated ratio the *runs test* was used (Freund and Simon 1992, 505-511). The runs test checks if there are either too many or too few runs of a particular category on a dichotomous scale. In this case the two categories were "less than the median" and "greater than or equal to the median". If there are a very small number of long runs of one category, then this indicates that consecutive trials are clustered or grouped. If there are very many short runs one might suspect some type of repeated alternating or cyclical pattern. Therefore to accept randomness one would look for a moderate number of runs of varying lengths. In the runs test the null hypothesis is that values of the dichotomous variable are randomly ordered; the alternate hypothesis is that values are not randomly ordered. Therefore to establish randomness one must fail to reject the null hypothesis. For this reason, for the runs tests performed on Experiments 1, 2 and 3, a liberal alpha-level of 0.1 was taken to reduce the probability of accepting the null hypothesis of randomness when runs of ratios were not randomly ordered.

Results

For Experiment 1, with the median value of 1.0010, there were 76 runs, 78 ratios less than the median, 82 ratios greater than or equal to the median, giving a total number of 160 ratios. From these numbers: $z = -.79$, $p = .43$ (two-tailed), therefore one cannot reject the null hypothesis that the ordering of ratios above and below the median ratio is random in Experiment 1.

For Experiment 2, with the median value of 1.0019, there were 59 runs, 64 ratios less than the median, 64 ratios greater than or equal to the median, giving a total of 128 ratios. From these numbers: $z = -1.06$, $p = .29$ (two-tailed), therefore one cannot reject the null hypothesis that the ordering of ratios above and below the median ratio is random in Experiment 2.

For Experiment 3, with the median value of .9950, there were 66 runs, 64 ratios less than the median, 64 ratios greater than or equal to the median, giving a total of 128 ratios. From these numbers: $z = .18$, $p = .86$ (two-tailed), therefore one cannot reject the null hypothesis that the ordering of ratios above and below the median ratio is random in Experiment 3.

In all three experiments, the ordering of ratios below and above the median displayed no systematic deviation according to the runs test.

In other words, deviations around the median were random over the time period of minutes, hours and weeks.

Conclusion

To conclude, then, one may say that the tape machine and the tapes used in all three experiments were very consistent. Over the period of four months, and with considerable use, the tape playing speed varied by less than +/-1%. Moreover, this minuscule deviation occurred randomly across trials, orders and measuring sessions. These results suggest that the experimental apparatus was suitable for use with stimuli for which consistency of IOI was of paramount experimental importance.

BIBLIOGRAPHY

Agmon, Eyton. 1990. Music theory as cognitive science: Some conceptual and methodological issues. *Music Perception* 7 (3): 285-308.

Allan, Lorraine G., and A.B. Kristofferson. 1974. Psychophysical theories of duration discrimination. *Perception and Psychophysics* 16 (1): 26-34.

Baddeley, Alan D. 1990. *Human Memory: Theory and Practice.* Boston: Allyn and Bacon.

Beaugrande, Robert de. 1985. General constraints on process models of language comprehension. In *Issues in Cognitive Modelling*, ed. A.M. Aitkenhead and J.M. Slack. London: Lawrence Erlbaum Associates.

Bengtsson, Ingmar, and Alf Gabrielsson. 1983. Analysis and synthesis of musical rhythm. In *Studies of Music Performance*, ed. Johann Sundberg. Stockholm: Royal Swedish Academy of Music.

Berry, Wallace. 1987. *Structural Functions in Music.* New York: Dover. (Originally published Englewood Cliffs, NJ.: Prentice Hall, 1976).

Berz, William L. 1995. Working memory in music: A theoretical model. *Music Perception* 12 (3): 353-64.

Blakely, W.A. 1933. The discrimination of short empty temporal intervals. Unpublished doctoral dissertation, University of Illinois, Urbana, IL.

Block, Richard A. 1990. Models of psychological time. In *Cognitive Models of Psychological Time*, ed. Richard A. Block. Hillsdale, NJ.: Lawrence Erlbaum.

Brittin, Ruth V. 1992. Perceptions of conducting: Accuracy in detecting modulated beat. *Bulletin of the Council for Research in Music Education* 113: 1-8.

Brittin, Ruth V. 1993. Discrimination of aural and visual tempo modulation. *Bulletin of the Council for Research in Music Education* 116: 23-32.

Brown, Peter. 1979. An enquiry into the origins and nature of tempo behaviour. *Psychology of Music* 7 (1): 19-35.

Brown, Peter. 1981. An enquiry into the origins and nature of tempo behaviour. Part II: Experimental work. *Psychology of Music* 9 (2): 32-43.

Bryden, M.P. 1971. Attentional strategies and short-term memory in dichotic listening. *Cognitive Psychology* 2 (1): 99-116.

Burns, Edward M., and W. Dixon Ward. 1978. Categorical perception--phenomenon or epiphenomenon: Evidence from experiments in the perception of melodic musical intervals. *Journal of the Acoustical Society of America* 63 (2): 456-68.

Clarke, Eric F. 1985. Structure and expression in rhythmic performance. In *Music Structure and Cognition*, ed. Peter Howell, Ian Cross and Robert West. London: Academic Press.

Clarke, Eric F. 1987a. Categorical rhythm perception: An ecological perspective. In *Action and Perception in Rhythm and Music*, ed. Alf Gabrielsson. Stockholm: Royal Swedish Academy of Music.

Clarke, Eric F. 1987b. Levels of structure in the organization of musical time. *Contemporary Music Review* 2: 211-38.

Clarke, Eric F. 1989. The perception of expressive timing in music. *Psychological Research* 51 (1): 2-9.

Cooper, Grosvenor W., and Leonard B. Meyer. 1960. *The Rhythmic Struture of Music*. Chicago: The University of Chicago Press.

Cowan, Nelson. 1984. On short and long auditory stores. *Psychological Bulletin* 96 (2): 341-70.

Cox, D.R., and D. Oakes, 1984. *Analysis of Survival Data*. London: Chapman and Hall.

Croonen, W.L.M., and P.F.M. Kop. 1989. Tonality, tonal scheme, and contour in delayed recognition of tone sequences. *Music Perception* 7 (1): 49-68.

Crowder, Robert G. 1982. Decay of auditory memory in vowel discrimination. *Journal of Experimental Psychology: Learning, Memory and Cognition* 8 (2): 153-63.

Crowder, Robert G. 1993. Auditory memory. In *Thinking in Sound: The Cognitive Psychology of Human Audition*, ed. Stephen McAdams and Emmanuel Bigand. Oxford: Oxford University Press.

Crowder, Robert G., and John Morton. 1969. Precategorical acoustic storage (PAS). *Perception and Psychophysics* 5 (6): 365-73.

Cuddy, Lola L., and H.I. Lyons. 1981. Musical pattern recognition: A comparison of listening to and studying tonal structures and tonal ambiguities. *Psychomusicology* 1 (2): 15-33.

Darwin, Christopher J., Michael T. Turvey, and Robert G. Crowder. 1972. An auditory storage analogue of the Sperling partial report procedure: Evidence for brief auditory storage. *Cognitive Psychology* 3 (2): 255-67.

Dawe, Lloyd A., John R. Platt, and Ronald J. Racine. 1993. Harmonic accents in inference of metrical structure and perception of rhythm patterns. *Perception and Psychophysics* 54 (6): 794-807.

Dewitt, Lucinda A., and Robert G. Crowder. 1986. Recognition of novel melodies after brief delays. *Music Perception* 3 (3): 259-74.

Dorhout, Albert Jensen. 1980. An investigation into the nature of tempo perception. *Dissertation Abstracts International* 40 (10): 5354A-55A. (Ph.D. Dissertation, Northwestern University, Evanston, IL., 1979).

Dowling, W. Jay, and James C. Bartlett. 1981. The importance of interval information in long-term memory for melodies. *Psychomusicology* 1 (1): 30-49.

Drake, Carolyn, W. Jay Dowling, and Caroline Palmer. 1991. Accent structures in the reproduction of simple tunes by children and adult pianists. *Music Perception* 8 (3): 315-34.

Duke, Robert A. 1989a. Effect of melodic rhythm on elementary students' and college undergraduates' perceptions of relative tempo. *Journal of Research in Music Education* 37 (4): 246-57.

Duke, Robert A. 1989b. Musicians' perception of beat in monotonic stimuli. *Journal of Research in Music Education* 37 (1): 61-71.

Duke, Robert A. 1994. When tempo changes rhythm: The effect of tempo on nonmusicians' perception of rhythm. *Journal of Research in Music Education* 42 (1): 27-35.

Duke, Robert A., John M. Geringer, and Clifford K. Madsen. 1991. Performance of perceived beat in relation to age and music training. *Journal of Research in Music Education* 39 (1): 35-45.

Efron, Robert. 1970. Effects of stimulus duration on perceptual onset and offset latencies. *Perception and Psychophysics* 8 (4): 231-34.

Ellis, Mark C. 1991. Research note. Thresholds for detecting tempo change. *Psychology of Music* 19 (2): 164-69.

Ellis, Mark C. 1992. Tempo perception and performance of elementary students, grades 3-6. *Journal of Research in Music Education* 40 (4): 329-41.

Erickson, Robert. 1975. *Sound Structure in Music.* Berkeley and Los Angeles: University of California Press.

Essens, Peter J. 1986. Hierarchical organization of temporal patterns. *Perception and Psychophysics* 40 (2): 69-73.

Fraisse, Paul. 1963. *Psychology of Time.* New York: Harper.

Fraisse, Paul. 1982. Rhythm and tempo. In *Psychology of Music*, ed. Diana Deutsch. New York: Academic Press.

Fraisse, Paul. 1984. Perception and estimation of time. *Annual Review of Psychology* 35: 1-36.

Fraisse, Paul, P. Pichot, and G. Clairouin. 1949. Les aptitudes rythmiques. Étude comparée des oligophrènes et des enfants normaux. *Journal de Psychologie Normale et Pathologique* 42: 309-330.

Freund, John E., and Gary A. Simon. 1992. *Modern Elementary Statistics.* 8th ed. Englewood Cliffs, NJ.: Prentice Hall.

Frick, Robert W. 1988. Issues of representation and limited capacity in the auditory short-term store. *British Journal of Psychology* 79 (2): 213-40.

Frischeisen-Köhler, Ida. 1933. The personal tempo and its inheritance. *Character and Personality* 1: 301-313.

Gabrielsson, Alf. 1973. Adjective ratings and dimension analyses of auditory rhythm patterns. *Scandinavian Journal of Psychology* 14 (2): 244-60.

Gabrielsson, Alf. 1974. Performance of rhythm patterns. *Scandinavian Journal of Psychology* 15 (1): 63-72.

Geringer, John M., Robert A. Duke, and Clifford K. Madsen. 1992. Musician's perception of beat note: Regions of beat change in modulating tempos. *Bulletin of the Council for Research in Music Education* 114: 21-33.

Geringer, John M., and Clifford K. Madsen. 1984. Pitch and tempo discrimination in recorded orchestral music among musicians and nonmusicians. *Journal of Research in Music Education* 32 (3): 195-204.

Getty, David J. 1975. Discrimination of short temporal intervals: A comparison of two models. *Perception and Psychophysics* 18 (1): 1-8.

Glucksberg, Sam, and George N. Cowen Jr. 1970. Memory for nonattended auditory material. *Cognitive Psychology* 1: 149-156.

Gordon, Edwin. 1976. *Tonal and Rhythmic Patterns: An Objective Analysis*. Albany, N.Y.: State University of New York Press.

Gross, Gerard, and Serge Daan. 1985. The use of biological clocks in time perception. In *Time, Mind, and Behavior*, ed. John A. Michon and Janet L. Jackson. Berlin: Springer-Verlag.

Guttman, Newman, and Bela Jullesz. 1963. Lower limits of auditory periodicity analysis. *Journal of the Acoustical Society of America* 35 (4): 610.

Halpern, Andrea R. 1988. Perceived and imagined tempos of familiar songs. *Music Perception* 6 (2): 193-202.

Halpern, Andrea R., and Christopher J. Darwin. 1982. Duration discrimination in a series of rhythmic events. *Perception and Psychophysics* 31 (1): 86-89.

Handel, Stephen. 1973. Temporal segmentation of repeating auditory patterns. *Journal of Experimental Psychology* 101 (1): 46-54.

Handel, Stephen. 1989. *Listening: An Introduction to the Perception of Auditory Events*. Cambridge, MA.: MIT Press.

Handel, Stephen, and Gregory R. Lawson. 1983. The contextual nature of rhythmic interpretation. *Perception and Psychophysics* 34 (2): 103-120.

Handel, Stephen, and James S. Oshinsky. 1981. The meter of syncopated auditory polyrhythms. *Perception and Psychophysics* 30 (1): 1-9.

Harrison, Ross. 1941. Personal tempo and the interrelationship of voluntary and maximal rates of movement. *Journal of General Psychology* 24: 343-379.

Hirsh, Ira J., Caroline B. Monahan, Ken W. Grant, and Punita G. Singh. 1990. Studies in auditory timing: 1. Simple patterns. *Perception and Psychophysics* 47 (3): 215-26.

Howell, Peter, and Christopher J. Darwin. 1977. Some properties of auditory memory for rapid formant transitions. *Memory and Cognition* 5 (6): 700-708.

Jackendoff, Ray. 1992. Musical processing and musical affect. *Cognitive Bases of Musical Communication*, ed. Mari Riess Jones and Susan Holleran. Washington, DC.: American Psychological Association.

James, William. 1950. *Principles of Psychology*. New York: Dover. (Originally published New York: Holt, Rinehart and Winston, 1890).

Jones, Mari Riess. 1976. Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review* 83 (5): 323-55.

Jones, Mari Riess. 1990. Musical events and models of musical time. In *Cognitive Models of Psychological Time*, ed. Richard A. Block. Hillsdale, NJ.: Lawrence Erlbaum Associates.

Jones, Mari Riess. 1992. Attending to musical events. In *Cognitive Bases of Musical Communication*, ed. Mari Riess Jones and Susan Holleran. Washington, DC.: American Psychological Association.

Jones, Mari Riess, and Marilyn Boltz. 1989. Dynamic attending and responses to time. *Psychological Review* 96 (3): 459-91.

Jones, Mari Riess, and William Yee. 1993. Attending to auditory events: The role of temporal organization. In *Thinking in Sound: The Cognitive Pyschology of Human Audition*, ed. Stephen McAdams and Emmanuel Bigand. Oxford: Clarendon Press.

Kalbfleisch, John D., and Ross L. Prentice. 1980. *The Statistical Analysis of Failure Time Data*. New York: John Wiley and Sons.

Kärnbach, Christian. 1994. Personal electronic-mail communication.

Keppel, Geoffrey. 1991. *Design and Analysis: A Researcher's Handbook*. 3rd ed. Englewood Cliffs, NJ.: Prentice Hall.

Kubovy, Michael, and Fred P. Howard. 1976. Persistence of a pitch-segregating echoic memory. *Journal of Experimental Psychology: Human Perception and Performance* 2 (4): 531-37.

Kuhn, Terry Lee. 1974. Discrimination of modulated beat tempo by professional musicians. *Journal of Research in Music Education* 22 (4): 270-77.

Kuhn, Terry Lee. 1987. The effect of tempo, meter, and melodic complexity on the perception of tempo. In *Applications of Research in Music Behavior*, ed. Clifford K. Madsen and C.A. Prickett. Tuscaloosa, AL.: University of Alabama Press.

Lapidaki, Eleni, and Peter R. Webster. 1991. Consistency of tempo judgments when listening to music of different styles. *Psychomusicology* 10 (1): 19-30.

Lachman, Roy, Janet L. Lachman, and Earl C. Butterfield. 1979. *Cognitive Psychology and Information Processing: An Introduction*. Hillsdale, NJ.: Lawrence Erlbaum Associates.

Large, Edward, J.D. McAuley, and J.F. Kolen. 1993. Adaptive oscillators for processing patterns in time. Paper presented at the Annual Conference of the Society for Music Perception and Cognition, Philadelphia, PA., June 1993.

Lee, Christopher S. 1985. The rhythmic interpretation of simple musical sequences: Towards a perceptual model. In *Music Structure and Cognition*, ed. Peter Howell, Ian Cross and Robert West. London: Academic Press.

Lee, Christopher S. 1991. The perception of metrical structure: Experimental evidence and a model. In *Representing Musical Structure*, ed. Peter Howell, Robert West and Ian Cross. London: Academic Press.

Lerdahl, Fred, and Ray Jackendoff. 1983. *A Generative Theory of Tonal Music*. Cambridge, MA.: MIT Press.

Longuet-Higgins, H. Christopher, and Christopher S. Lee. 1982. The perception of musical rhythms. *Perception* 11: 115-128.

Longuet-Higgins, H. Christopher, and Christopher S. Lee. 1984. The rhythmic interpretation of monophonic music. *Music Perception* 1 (4): 424-41.

Lu, Z.-L., S.J. Williamson, and L. Kaufman. 1992. Behavioral lifetime of human auditory sensory memory predicted by physiological measures. *Science* 258: 1668-70.

Luce, R. Duncan. 1986. *Response Times: Their Role in Inferring Elementary Mental Organization*. Oxford Psychology Series No. 8. New York: Oxford University Press.

Mabbott, J.D. 1951. Our direct experience of time. *Mind* 60: 153-67.

Madsen, Clifford K. 1979. Modulated beat discrimination among musicians and non-musicians. *Journal of Research in Music Education* 27 (2): 57-67.

Madsen, Clifford K., Robert A. Duke, and John M. Geringer. 1984. Pitch and tempo discrimination in recorded band music among wind and percussion musicians. *Journal of Band Research* 20 (1): 20-29.

McClave, J.T., and F.H. Dietrich. 1991. *Statistics*. 5th ed. Toronto: Collier Macmillan.

Michon, John A. 1978. The making of the present: A tutorial review. In *Attention and Performance* VII, ed. Jean Requin. Hillsdale, NJ.: Lawrence Erlbaum.

Michon, John A. 1985. The compleat time experiencer. In *Time, Mind, and Behavior*, ed. John A. Michon and Janet L. Jackson. Berlin: Springer-Verlag.

Miller, George A. 1956. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review* 63 (2): 81-97.

Moore, Brian C.J. 1982. *An Introduction to the Psychology of Hearing.* 2nd ed. London: Academic Press.

Morton, John, Robert G. Crowder, and Harvey A. Prussin. 1971. Experiments with the stimulus suffix effect. *Journal of Experimental Psychology*, 91 (1): 169-90.

Norušis, Marija J. 1991. *The SPSS Guide to Data Analysis for SPSS/PC+, 2nd Edition*. Chicago, IL.: SPSS Inc.

Neisser, Ulric. 1967. *Cognitive Psychology.* New York: Appleton-Century-Crofts.

Pachella, Robert G. 1974. The interpretation of reaction time in information-processing research. In *Human Information Processing: Tutorials in Performance and Cognition*, ed. Barry H. Kantowitz. Hillsdale, NJ.: Lawrence Erlbaum Associates.

Parncutt, Richard. 1994. A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception* 11 (4): 409-464.

Pöppel, Ernst. 1972. Oscillations as possible basis for time perception. In *The Study of Time*, ed. J.T. Fraser, F.C. Haber and G.H. Müller. Berlin: Springer-Verlag.

Povel, Dirk-Jan. 1981. Internal representation of simple temporal patterns. *Journal of Experimental Psychology: Human Perception and Performance* 7 (1): 3-18.

Povel, Dirk-Jan. 1984. A theoretical framework for rhythm perception. *Psychological Research* 45 (4): 315-37.

Povel, Dirk-Jan, and Peter Essens. 1985. Perception of temporal patterns. *Music Perception* 2 (4): 411-40.

Povel, Dirk-Jan, and Hans Okkerman. 1981. Accents in equitone sequences. *Perception and Psychophysics* 30 (6): 565-72.

Rimoldi, Horacio J.A. 1951. Personal tempo. *Journal of Abnormal and Social Psychology* 46 (3): 280-303.

Rostron, A.B. 1974. Brief auditory storage: Some further observations. *Acta Psychologica* 38 (6): 471-82.

Royal, Matthew S. 1993. The perception of tempo modulation in drum rhythms. Unpublished paper presented at the Regional Graduate Music Student Colloquium, University of Toronto (October 30, 1993).

Sanford, Anthony J. 1985. *Cognition and Cognitive Psychology*. London: Weidenfeld and Nicolson.

Schulze, Hans-Henning. 1989a. Categorical perception of rhythmic patterns. *Psychological Research* 51 (1): 10-15.

Schulze, Hans-Henning. 1989b. The perception of temporal deviations in isochronous patterns. *Perception and Psychophysics* 45 (4): 291-96.

Seashore, Carl E. 1967. *Psychology of Music*. New York: Dover. (Originally published New York: McGraw-Hill, 1938).

Shaffer, L. Henry. 1981. Performances of Chopin, Bach and Bartok: Studies in motor programming. *Cognitive Psychology* 13: 326-76.

Siegel, Jane A., and William Siegel. 1977. Categorical identification of tonal intervals: Musicians can't tell *sharp* from *flat*. *Perception and Psychophysics*, 21 (5): 399-407.

Sloboda, John A. 1983. The communication of musical metre in piano performance. *Quarterly Journal of Experimental Psychology* 35A: 377-96.

Sloboda, John A. 1985. Expressive skill in two pianists: Metrical communication in real and simulated performances. *Canadian Journal of Psychology* 39 (2): 273-93.

Smith, Jacqui. 1983. Reproduction and representation of musical rhythms: The effects of musical skill. In *Acquisition of Symbolic Skills*, ed. David Rogers and John A. Sloboda. New York: Plenum Press.

Smith, Karen C., and Lola L. Cuddy. 1989. Effects of metric and harmonic rhythm on the detection of pitch alterations in melodic sequences. *Journal of Experimental Psychology: Human Perception and Performance* 15 (3): 457-71.

Stevens, Jospeh C., and James W. Hall. 1966. Brightness and loudness as functions of stimulus duration. *Perception and Psychophysics* 1: 319-27.

Stott, L.H. 1933. The discrimination of short tonal durations. Unpublished doctoral dissertation, University of Illinois, Urbana IL.

Tabachnick, Barbara G., and Linda S. Fidell. 1989. *Using Multivariate Statistics*. 2nd ed. New York: Harper and Row.

Thomassen, Joseph M. 1982. Melodic accent: Experiments and a tentative model. *Journal of the Acoustical Society of America* 71 (6): 1596-1605.

Todd, Neil. 1985. A model of expressive timing in music. *Music Perception* 3 (1): 33-58.

Treisman, Anne. 1964. Monitoring and storage of irrelevant messages in selective attention. *Journal of Verbal Learning and Verbal Behavior* 3: 449-59.

Treisman, Michel, and A.B. Rostron. 1972. Brief auditory storage: A modification of Sperling's paradigm. *Acta Psychologica* 36 (2): 161-70.

Vos, Joos, and Rudolf Rasch. 1981. The perceptual onset of musical tones. *Perception and Psychophysics* 29 (4): 323-35.

Wallin, J.E. Wallace. 1911. Experimental studies of rhythm and time: The preferred length of interval (tempo). *Psychological Review* 18 (3): 202-22.

Wang, Celia Chu. 1983. Discrimination of modulated music tempo by music majors. *Journal of Research in Music Education* 31 (1): 49-55.

Wang, Celia Chu. 1984. Effects of some aspects of rhythm on tempo perception. *Journal of Research in Music Education* 32 (3): 169-76.

Wang, Celia Chu, and Rita S. Salzberg. 1984. Discrimination of modulated music tempo by string students. *Journal of Research in Music Education* 32 (2): 123-31.

Wapnick, Joel. 1980. The perception of musical and metronomic tempo change in musicians. *Psychology of Music* 8 (1): 3-12.

Welford, A.T. 1980. Choice reaction time: Basic concepts. In *Reaction Times*, ed. A.T. Welford. London: Academic Press.

Wickelgren, Wayne A. 1969. Associative strength theory of recognition memory for pitch. *Journal of Mathematical Psychology* 6 (1): 13-61.

Windsor, W. Luke. 1993. Dynamic accents and the categorical perception of metre. *Psychology of Music* 21 (2): 127-40.

Woodrow, Herbert. 1951. Time perception. In *Handbook of Experimental Psychology*, ed. S.S. Stevens. New York: Wiley.

Wundt, Wilhelm. 1911. *Grundzüge der physiologischen Psychologie.* Leipzig: Wilhelm Engelmann.

Yarbrough, Cornelia. 1987. The effect of musical excerpts on tempo discriminations and preferences of musicians and non-musicians. In *Applications of Research in Music Behavior*, ed. Clifford K. Madsen and C.A. Prickett. Tuscaloosa, AL.: University of Alabama Press.

Yeston, Maury. 1976. *The Stratification of Musical Rhythm*. New Haven, CN.: Yale University Press.

Zwislocki, J.J. 1972. A theory of central auditory masking and its partial validation. *Journal of the Acoustical Society of America* 52 (2, part 2): 644-59.