METHUEN'S
MANUALS OF MODERN PSYCHOLOGY

Edited by C. A. Mace

The Psychology of Musical Ability

The Psychology of Musical Ability

ROSAMUND SHUTER
Lecturer in Psychology,
West Ham College of Technology

METHUEN & CO LTD
11 New Fetter Lane London EC4

First published in 1968

Printed in Great Britain by
Butler & Tanner Ltd
Frome and London

Distributed in the USA
by Barnes & Noble Inc.

To Dr Herbert Wing
with respect and affection

Contents

PART IV

Theories of Musical Ability

PART V

Music and Education

Foreword

There is a long tradition of musical psychologists, or at least of psychologists who have devoted much of their research efforts to exploring the problems raised by music and its effects on people. Von Helmholtz first published his *Sensations of Tone* just over 100 years ago. He was followed – to select but a few names – by Stumpf and Revesz in Europe, Seashore, Farnsworth, Mursell and many others in the United States, and in Britain by Myers, Valentine and Wing. Music poses particularly fascinating problems, not merely as an art form and a source of intense interest and enjoyment to a great number of people. It provides a series of stimuli whose physical constitution is simpler and better understood than that of vocal or other auditory stimuli; and by now we know a good deal about how the ear picks up these stimuli and conveys them to the brain. But there is a tremendous amount to find out about the way in which we come to perceive these arbitrary combinations of sound waves as a kind of language *sui generis*, composed of melodies, harmonies and rhythms, and how these elicit the whole gamut of human emotions. Why, again, does music vary so widely in its effects from one person to another, from the 'droner' who is incapable of singing in tune to the infant prodigy like Mozart?

It is this aspect which constitutes Mrs Shuter's theme, and she is, I believe, the first to publish a book on this subject since Seashore's *Psychology of Musical Talent* in 1919. The advances that she has to record over 48 years are indeed astonishing. She has dealt exhaustively with a very wide range of research on musical abilities, how they can be recognised and tested, how far they should be attributed to heredity or to factors of upbringing and training, how are they inter-related and connected with intelligence and other abilities. As she herself points out, this is 'a progress report' on what has been learnt so far, not a series of final conclusions, though she is able to draw attention to many important educational and other applications. Among her major arguments which, I would agree, are substantiated by the evidence she presents, are that heredity is indeed important in making some children more sensitive to musical stimuli than others and in setting a limit to an individual's musical capabilities. But at the same time much more could be done than usually is done at present to stimulate and develop sensitivity, appreciation and vocal and instrumental performance, among the great bulk of the population. For example the 'droner' is by no means hopeless. The recent introduction of programmed learning and other new educational techniques offers scope for a vast expansion in the effective teaching of music. We already have tests which, though they show various limitations, are capable of picking out the talented pupil who deserves special encouragement by the age of 11 or earlier. Indeed I would go so far myself as to claim that we can probably make better predictions from childhood to adult accomplishment in music than in any other specialist field. However this should not, of course, be taken to imply that talent and sensitivity always show themselves early, or that people who display little musical inclination in childhood cannot develop good appreciation or become useful performers during adolescence and adulthood.

Much of the research that Mrs Shuter describes is inevitably technical; but she has gone a long way to make the concepts and findings intelligible to the layman, and to organise the material lucidly. Thus while this survey of research will be invaluable to the psychologist, it also has much of interest to offer to the educationist, especially the music teacher, and to the musically inclined parent.

PHILIP E. VERNON

Introduction

The common man of today is far richer in music than the prince of two centuries ago. Through radio and television he can command the services of not just one orchestra, but of many. Though he may not be able to commission an ode in celebration of his birthday, he can, for the price of a record, have the finest artists performing in his own home any music that happens to suit his mood at the moment. Yet relatively few people take full advantage of this wealth. Many who do try to listen attentively are often only too well aware how much they are missing of the real meaning of the music. Almost everyone understands his native language, but many people cannot sing in tune, much less play an instrument. Most people learn to read; fewer are musically literate. We may wonder whether this is due to a lack of interest or of opportunity to learn. Or is it necessary, in order to understand and appreciate music, to have some special gift?

What in fact makes a musician? What distinguishes the musician from the rest of mankind? Why should one infant in the cradle seem so much more responsive to music than another? Why should one, but only one, of the small town German musicians at the beginning of the seventeenth century have been the progenitor of the amazing Bach family who produced in six generations no less than 47 musicians of talent or genius? Was this due to heredity or to family background? Is there a sense in which we are all musical? How can we educate our children to enjoy and truly appreciate a wide variety of music?

It is with questions such as these that this book is primarily concerned. It evolved from a doctoral thesis and from the writer's ponderings on the barrier to advanced achievement in music which she had earlier encountered and which, in spite of her own and her teachers' efforts, had proved insurmountable. The writer's aim has been to collate and evaluate psychological studies of musical ability and attainment in order to make the results more widely accessible to all who are interested in music, whether as performers, teachers or listeners. Much less research has been carried out into musical ability than into other abilities, e.g. 'intelligence'. However, though fragmentary, studies of some aspects of musical ability have produced interesting and important results. The present book is not concerned, except incidentally, with the psychology of sound, nor with the aesthetic aspects except in relation to ability to compose or appreciate music, nor with music therapy which is fast becoming a specialist area.

Historical Background

Interest in the psychology of music is as old as experimental psychology itself, the birth of which may be dated at 1879 when Wilhelm Wundt opened his 'Psychological Institute' at the University of Leipzig. Wundt and many of his associates were physiologists by training. They concentrated at first on measurements of sensitivity to auditory, and other sensory stimuli, and of simple reaction time. Gradually, however, their attention turned to the measurement of perceptual span and the rate of learning. The need for rigorous control of experimental conditions soon became apparent; for it was found that even such factors as the wording of instructions could significantly influence the results obtained. The acoustical studies of the pioneers of experimental psychology such as Hermann von Helmholtz and Carl Stumpf (both professors in the University of Berlin) are discussed in Geza Revesz's *Introduction to the Psychology of Music.*

Stumpf was himself a cellist and as early as 1883 had already devoted considerable thought to possible tests of musical aptitude. He had become interested in individual differences in ability after observing the varying responses obtained with musical and unmusical persons during his study of tones. He carried out experimental tests with David Popper, a celebrated cellist, and with Pepito Areola, an infant prodigy. Revesz was in many ways Stumpf's direct successor. He too studied – for five years – an infant prodigy and experimented with tests which closely resembled those given traditionally by music teachers. He left his native Hungary after the First World War and settled in Holland. Though greatly influenced by the Gestalt school of psychology then prominent in Europe, his own theories were too wide to be easily classified. Among the many investigations he carried out was a study of the popular notion that there is a connection between mathematical ability and a talent for music.

On the theoretical side, the first important book on the psychology of musical ability *Wer ist musikalisch?* appeared in 1895. The author was Theodor Billroth, a Vienna physiologist and music lover. The characteristics of the musical person which he listed are of a relatively objective nature. He recognised the importance of a spontaneous interest in music and memory for musical material, as well as the sensory capacities he had studied as a physiologist. Another physiologist with musical training, Johannes von Kries, wrote in 1926 a book with a similar title to Billroth's in which he attempted to collate all the characteristics of the musical person.

The value of the research into problems of the psychology of musical ability at this period was limited by the lack of controlled tests. The inheritance of musical talent was studied by Francis Galton's methods of drawing up family trees on the basis of reputation or – a later refinement – of answers to questionnaires. Galton himself realised the importance of trying to make the most accurate possible assessment of individual abilities and of finding statistical methods of comparing different groups of people and the results of different tests. Among his many inventions was the Galton whistle for determining the highest audible pitch. He believed that powers of sensory discrimination were directly related to intellectual powers. It was not long before he was proved wrong as far as general intellectual ability was concerned.

Meanwhile in the United States, the full resources of the psychological laboratory were being brought to bear on the problems of providing objective tests of musical ability. Around 1890 W. E. Scripture set up at Yale University a psychological laboratory on the lines of Wundt's at Leipzig. Numerous experiments were carried out on vision, hearing and the other senses. It was here that the first measurements of pitch discrimination were made by a group test. Here too Hughes was the first to compare the scores obtained under laboratory conditions with an outside criterion of musical ability.

But the most important of the pioneers at Yale, from the point of view of music, was Carl Emil Seashore. In 1897 he went to the State University of Iowa where he stayed for 40 years and later became director of the Psychology Laboratory. His pioneer work included the invention of the Voice Tonoscope, which gives a visual picture of a tone, thus enabling the singer to see the sound he is producing, and the audiometer, an instrument for measuring the threshold of hearing for the intensity of sounds at various frequencies. In his books *The Psychology of Musical Talent* (1919) and *The Psychology of Music* (1938) Seashore emphasised that his aim was to apply the technological apparatus and methods of his laboratory for *the service of music*. His Measures of Musical Talent, published in 1919 after twenty years of experimentation were intended to select for training gifted children whose talents might otherwise be neglected and to save the unmusical and their teachers from the discouragement of failing to progress. The publication of his measures was followed by many efforts to check their validity in actual prognostic situations. Criticisms were directed at Seashore on several accounts: James Mursell of Columbia University Teachers' College doubted whether measurements of isolated, specific capacities could have much relevance to functional musical activities. Paul Farnsworth of Stanford and Robert Lundin then of Hamilton College, New York, besides producing evidence of the deficiencies of the Measures, also attacked Seashore for his assumption that the capacities he was seeking to assess were innate and unaffected by training. Seashore himself was well aware of the more complex aspects of musical perception and performance but believed that the capacities he was testing were as basic to musical aptitude as they were to sound itself. Paradoxically, this author of the most objective, laboratory-bound tests was prone to make *ex cathedra* pronouncements which seem, in these less confident days, not wholly in the spirit of scientific enquiry. He was not himself a musician, though he is said to have played the organ when he was young. But the sincerity of his interest in music cannot be doubted. Many who have criticised his approach to testing musical ability have themselves adopted his aims.

The criticisms of Seashore suggested that tests based on musical material might prove more satisfactory provided that they could be properly standardised. During the 1930s three successful batteries were in fact developed. Kate Hevner, the only woman author of a test that has gained an international reputation, produced the Oregon Music Discrimination test. Raleigh Drake, a musician as well as a psychologist, devised during a course of postgraduate study in London, a musical memory test of lasting worth. Herbert Wing, who had practical experience of schoolteaching as well as being a musician and psychologist, developed a comprehensive battery of aural acuity and musical appreciation tests. His research at University College, London was guided by Cyril Burt who included music in his immensely wide interests. Burt had concluded from his early experiments with the Seashore tests that they were unsatisfactory.

Since 1940 other tests based on musical material have been produced. Validation studies of these and the earlier tests have been continued and tests have gradually come to enjoy increasing use both in educational situations and as research tools. More sophisticated forms of statistical analysis have been applied to the results obtained with the tests to throw light both on exactly what the tests are measuring and on the complex nature of musical ability.

Tests have not been the only source of data for the psychologist of musical ability. Intensive case studies have also been made, concerned for example with the early development of individual children, or with the exceptional displays of talent of the musical prodigy, or of the idiot savant. Psychological experiments have also been carried out where one 'variable' is selected for study, all the other factors in the situation being as far as possible controlled. For example, many experiments have been devised to try and determine the effect of a particular type of training on the singing or performing ability of an experimental group of 'subjects'. The results on some standard task before and after the training is compared with those of a 'control' group who have not received the training.

Observations, both of spontaneous behaviour and of responses induced by the experimental conditions, can be a valuable source of data. Best results are obtained when the behaviour to be studied has been specified beforehand and where more than one observer works independently.

'Musical ability' is the term generally adopted throughout this book, as being in Farnsworth's words 'the broadest and safest', since it suggests power to perceive and act without any *a priori* implication to the extent of heredity. (Farnsworth, however, believes that we should speak of musical abilities.) 'Talent' has been used similarly, but usually with the implication of some positive degree of ability. 'Musical' is taken to mean simply 'having musical ability'. The reservation 'musical ability as assessed by the such and such test' frequently needs to be read into the text. Other writers on the psychology of musical ability have preferred other terms. Thus Holmstrom uses 'musicality' and Teplov, 'sens musical', while both musicians and laymen speak of 'an ear for music'. The term 'musical ear' should, of course, include not only the sensory and perceptual system, but also the integrating and interpreting power of the human mind.

But the possession of a fine ear, important prerequisite though it is, does not make a musician. To perform or sing the motor mechanism of the mind and body must be brought into play. Therefore in practice a useful distinction can be drawn between the perceptual and the muscular aspects.

To speak only of ability to perceive and understand music or to play it with nimble fingers would be to leave out of account the heart of the matter. For music is an art, music is beautiful. Its power to move, to excite and to charm has always been recognised. As Curt Sachs tells us in his fascinating essay 'The Lore of Non-Western Music', all over the ancient world from Egypt to China, beautiful women with musical training were a typical gift to royal friends or suzerains.

The contrast between the ability, if such it can be called, to respond emotionally to music and what we mean by musical ability is epitomised in two characters described by Stendhal in his *Life of Rossini*. One, an elderly clerk from the War Office, possessed the gift of absolute pitch to such a degree that if he happened to hear a couple of workmen on a building site chipping a block of stone with their hammers, he could tell instantly the exact notes which the two sets of hammer blows were emitting. He could also copy down any tune he heard. However, music apparently gave him no pleasure whatsoever, as an art it was meaningless to him. The other, a young Venetian nobleman, was incapable of singing four notes on end without committing the most excruciating cacophony. Yet he adored music with a passionate intensity, rare even in Italy.

Such an extreme contrast is in reality quite unusual. More commonly, emotional and intellectual elements combine in the highest type of aesthetic experience. But this is a matter which will be discussed at some length in Part Four of the book.

First, in Part One, we shall discuss how musicians and psychologists have tried to find ways of assessing musical ability and achievement that would prove useful in education and in research. Part Two deals with the development of musical ability from birth to adolescence, Part Three with hereditary and environmental factors in musical ability. Finally, in Part Five, some research findings relevant to music education are discussed. Though as much of the factual and statistical data as possible have been removed to the appendixes, the non-psychologist reader may prefer to regard for example Chapters II and III as for reference rather than continuous reading. Some conclusions on the topics dealt with in each part of the book are briefly summarised and discussed in the last chapter of each part.

PART I

The Assessment of Musical Ability and Attainment

I

General Problems of Testing

The Professor struck a chord on the piano. 'What chord is this?' he asked. 'It's a major chord, sir,' replied the candidate. 'What else can you tell me about it?' After a pause in which the candidate could not elaborate on his answer, the Professor sounded a discord. 'That's a major – no, minor seventh . . .' 'Try again,' said the Professor. But, though he played three more chords, he had already decided to reject the candidate. The second candidate was more fortunate; he possessed absolute pitch and had no difficulty in identifying the chords.

By such an arbitrary method were the candidates for places in the music department of a senior British university selected not so very long ago.

It does not take a psychologist to point out the deficiencies of such a procedure. The test was very brief and confined to only one type of item. With so much depending on their success, the candidates were likely to be nervous and had no time to settle. It is probable that a chord test had considerable discriminatory value with candidates who must have already passed through other screening examinations at earlier stages in their musical training. The successful candidates were no doubt found to be quite satisfactory by the Professor, who after all had selected them himself. But he had no means of knowing how well some of the rejected candidates might have succeeded in the course.

Music teachers and examiners usually employ much more rational procedures. In fact, Wing (1948, p. 5) remarked that the organisation of a well-conducted musical competition is as close an approximation to a psychological test as it could be hoped to obtain in an aesthetic activity. The conditions are made as standard as possible, with the assessment being done on a scale of marks. However, the music festival adjudicator falls short of the demands of the psychologist in that his marks tend to vary with the standard of the competition. In addition to standard tasks and an objective scoring procedure, the psychological test includes tables of the scores made by representative groups, so that it is possible to compare the testee with others of similar age.

It is no easy matter, as Lowery (1929) pointed out, to devise tests of musical ability, 'since a passage of music involves numerous factors which, in general, are not readily isolated from one to another; so that the experimenter who would have his subjects attend to the variation of some one factor in a series of presented phrases is often at a loss how to obtain phrases in which the special factors to which special attention is to be given may be pointed out quite unambiguously'. Lowery (1952) elsewhere relates that, when he first tried to formulate a cadence test on 'giving the test to both children and adults, chaotic results were obtained in spite of careful efforts to ensure the subjects understood what was required'.

Most of the tests so far developed have for practical reasons been intended for group application. Group tests enable results from the large numbers necessary for standardisation to be collected within a reasonably short time. They also enable the user to classify one or more classes of pupils at one session. Individual testing is, however, generally more reliable. This is partly due to the nature of group tests themselves, in so far as they usually have a number of multiple choice answers into which chance factors may enter. It is also due to the conditions of application; for if a child becomes confused or misunderstands the instructions, he may get a whole section wrong in a group test, whereas he could be put right in an individual test by the examiner. If some important decision hinges on the results of a group test, it is advisable to supplement it by an individual test and by such other evidence as is available, e.g. from teachers' reports.

Readers who are interested in the task of developing a test battery will find a detailed account of the evolution of Wing's harmony test in his monograph (1948). Briefly, the procedure is as follows: 'after deciding which aspects of musical ability he wishes to assess, the test author selects a number of possible items which seem suitable for his purpose. He then has to try them out on a variety of children, or adults, or both, to see which are the most satisfactory.

In the early stages it may not be possible to judge whether any unsatisfactory results are due to the music used, or to the method of application, or whether they are inherent in the kind of test.

Each item of each subtest has to be carefully examined to see whether it is contributing its share to the total marks. This is done by item analysis which shows how many subjects obtain the correct answer to each item and also whether those who get the right answers are the subjects who are known to be musical. After item analysis, unsatisfactory questions are revised or discarded.

When the test is in reasonably good shape after these preliminary trials, it is preferable to record it to avoid inevitable differences in playing.

Work can now begin on compiling the norms, i.e. the normal mark that would be obtained by a truly average child of a specified age. The groups whose scores are going to be used for the norms against which, when published, the results of testing will be compared, must be as representative as possible of the whole population on which the test is likely to be used. Subjects of both sexes, from all socio-economic classes, and from all parts of the country should be included.

Tests should preferably have been standardised in the country in which they are to be used, as norms may not always apply in other countries without some readjustment, despite the 'universal language' of music. The norms can be presented either as percentiles or as grades. With percentiles the score of the individual is interpreted in terms of what percentage of other individuals score above (or below) his score. For example, if only 20% of the group on which the test was standardised exceed his score, he is said to have a percentile rank (PR) of 80. PR 50 is the median or middle point of a range of scores. In other tests, the scores are divided into grades, e.g. the top 10% may be called grade 'A', etc. Wing, besides using grades, provides a formula by which 'Musical Age' can be calculated, on the analogy of 'Mental Age' in intelligence. The child of 10 whose score equals the average obtained by a 12-year-old child is said to have a Musical Age of 12, or to be two years in advance of his chronological age. By comparing musical with chronological age, Musical Quotients similar to the IQ can be worked out. Thus, a child aged 10, with a Musical Age of 12, would have a MQ of 120; if his Musical Age was only 8, his MQ would be 80.

In America, grade levels are often used instead of age levels. Since the age of entrance to school is six, grade 1 would consist of six-year-olds, but later grades might have more heterogeneous ages, as some of the children might be repeating a grade, while others were promoted more quickly than the general stream.

Some tests are primarily intended to give an index of general musical level, derived from total scores. Others provide a profile of the abilities measured by the subtests. Total scores give the more reliable results. Even when norms are provided only for the test total, some rough idea of performance of the various subtests can be gathered from inspecting the marks obtained for the individual tests.

As well as providing tables of norms, a test author can be expected to publish evidence of the *reliability* and *validity* of his tests.

Reliability refers to the consistency of a test in yielding the same, or closely comparable, results if given to the same subjects on subsequent occasions. It is shown statistically as a coefficient of correlation.

A coefficient of correlation is a measure of the degree of resemblance of two sets of scores or of two orders of merit. When calculated directly from scores, the coefficient is indicated by the symbol r, when rank orders are used, by ρ (rho). If the two sets of scores corresponded exactly, the coefficient would be 1·00. A perfect correlation rarely occurs in practice.

A coefficient of ·80 or ·90 is a very high correlation, such as is obtained by duplicate sets of a good intelligence test. A coefficient of ·40 or less indicates only a slight resemblance between the two sets of scores and is quite consistent with many considerable differences in the position of individuals. A resemblance between the two sets of scores no greater than would happen (on the average) by mere chance would be represented by ·00. If the two sets of figures are exactly the opposite to one another the figure would be −1; but these negative correlations do not occur often when comparing results from the same test on different occasions or even from different cognitive tests.

In interpreting correlation coefficients, the range of abilities of the group studied must be kept in mind. A correlation will be lower for a very selected group than for one which has a wide range of abilities. However, as a rough guide to the interpretation of reliability coefficients Leonhard and House (see Whybrew, 1962) have summarised opinions concerning usual degrees of reliability in the following table:

·85–·99	High to very high; of value for individual measurement and diagnosis.
·80–·84	Fairly high; of some value in individual measurement and highly satisfactory for group measurement.
·70–·79	Rather low; adequate for group measurement but of doubtful value in individual measurement.
·50–·69	Low; inadequate for individual measurement but of some value in group measurement.
Below ·50	Very low; inadequate for use.

The reliability of a test depends on a number of factors. One is length. The longer the test, the more likely it is to be reliable, at least till the point is reached where scores may be affected by fatigue or boredom.

Another important factor is the suitability of the test for the people on which it is being used. If there are, for example, too many difficult items, a large proportion of the answers may be based on guesswork. With a test which may be used with subjects of different ages and ability levels, it would be desirable to have reliability coefficients from a variety of populations. The reliability of a test can be assessed by giving it to the same group of people on two different occasions and correlating the results. This is the *test retest* method. Alternatively, two equivalent forms of the test might be given to the same group. A practical modification of this method is to compare odd and even items of one testing. This is known as the *split-half* method. A correction formula is applied to the result to compensate for the reduction by half of the number of items.

Another technique for estimating reliability was devised by Kuder and Richardson (see Whybrew, 1962). Based on an analysis of the subject's performance on each item, it in fact provides a measure of consistency between the items.

A test is said to be valid in so far as it measures what it purports to measure. Thus a test that claims to measure musical aptitude is valid only so far as it measures that and not intelligence, for example, or some other trait. One means of validating a test is by item analysis, as mentioned above. This ensures that the test is internally consistent and that each item is measuring what all the other items are measuring. Other methods involve comparing test scores with some outside criterion of musical ability. This may be teachers' ratings, examination marks for music or success in music as a profession.

It is not usually easy to obtain reliable ratings from class teachers, especially in the case of a subject like music. A teacher probably knows the bright pupils and the very weak ones, but may find it very difficult to rate all the members of a class. Teachers of specialist music classes or of instrumental pupils are familiar with the work of the individual pupils, but these are usually selected and not representative of the general population. Moreover, the numbers available are small. However careful and unbiased the rater tries to be his judgments may be subjective and inaccurate. When a low correlation between test and rating is obtained, it is sometimes difficult to decide whether this is due to deficiencies in the test, or to the unreliability of the rating, or to both.

Another means of validating a test is suggested in Mursell's remark – 'We must try our developed tests upon individuals known to be conspicuously musical and those known to be conspicuously nonmusical to try to discover where the most crucial and significant performances are located' (1937). If a music test has any validity at all, it must discriminate between recognised musicians and persons of average or low ability. A test that can discriminate not only between these, but also between the more and the less able members of a highly talented group can be considered to have superior validity.

The results of comparing groups should be tested to see whether the difference obtained is actually 'statistically significant'. Significance in connection with differences of scores does not mean 'worth noting'. It means that the likelihood of such a difference arising by mere chance is so slight that it is not worth considering. If a difference is reported to be highly significant, the likelihood of its being due merely to chance is 1 in 100.

The statistical procedure of factor analysis is sometimes adopted to isolate the significant variables of musical ability and to validate the appropriate tests. Factor analysis is a means of resolving a set of intercorrelating tests into a few factors which are regarded as being the fundamental underlying variables (see further, Appendix II). If a factor had high loadings on a number of pitch discrimination tests, but only weak loadings on rhythmic tests, it would be reasonable to consider it a pitch factor. The validity of any pitch test which did *not* have a high loading in this factor would be suspect.

A further method of validating a new test is by comparing it with an existing established test. This has become possible only comparatively recently in the case of musical ability tests, owing to the lack of previous tests of proven validity.

Tests can be broadly divided into two types: those of attainment where the aim is to assess what has been learned, and those of aptitude or potentiality which seek to predict future success. In music, attainment tests may take the form of questionnaires on musical knowledge, or of scales against which vocal or instrumental performance can be compared. Most prognostic tests so far developed deal only with the aural side; the motor skills required for musical performance have been but little investigated. It has usually been the aim to try to devise tests that are as little affected by previous experience of music as possible. However, some tests, though mainly aural, require a knowledge of notation. In addition, efforts have been made to find means of assessing interest in music, though these are perhaps not tests in the usual sense of the word.

II

Tests of Musical Ability

This chapter is divided into three parts. The first deals with earlier attempts to test musical ability that were not commercially published. This may have been because the author was primarily interested in research or because he was unable to spare the time to produce a standardised version. In the second section we shall discuss the more important of the tests which are commercially available. Finally, mention will be made of more recent attempts to develop tests, which have not so far been fully standardised.

Detailed description of tests which are commercially available, or which have at least enjoyed some use for research or educational purposes, can be found in Appendix I.

EARLIER UNSTANDARDISED TESTS

In the 1880s Stumpf devoted considerable study to possible tests of musical ability and devised a few simple tests that are very similar to those traditionally given by music teachers: singing a given note that had been struck on the piano; judging which was the higher of two notes played successively; and judging degrees of consonance for pleasantness. These were successful in discriminating between experienced musicians and 14 self-confessed 'unmusical' students.

In 1920 Revesz produced a more extensive battery of tests which like Stumpf's, required individual application. For example the subject was asked to imitate by clapping rhythmic patterns played on the piano or to sing the notes of chords that had been played. He also attempted to test 'regional pitch', a sort of approximate absolute pitch. Eight notes between G_1 A^3 were played on the piano in irregular order, the subject being asked to find each note on the piano. A test which Revesz believed to be particularly important was singing back melodies. He played nine bars of a tune, then repeated the first two bars, the subject being required to continue the melody. Revesz used the scores on this test as a criterion with which he correlated all the other tests. Playing from ear correlated ·77 and the pitch tests about ·60 with this criterion in his experiments with children aged 7–12.

In his book published in German in 1946, Revesz recommended the use of his tests of rhythm, regional pitch, two-note chords and ability to grasp and sing a tune as measures of the 'lower grades of musicality'. For the 'higher grades of musicality' he would give tests of relative pitch, harmonic apprehension and response, playing familiar tunes by ear and creative fantasy (singing the ending of a familiar, unfinished melody).

Franklin considered the two rhythmic tests to be among the best in Revesz's battery and used them in a modified form for his own investigation (see p. 301). His results largely verified Revesz's.

Revesz did not standardise his tests and did not intend them as group tests. If, as Drake (1931) suggested, norms could be established for them, they might prove very useful to private music teachers.

Schoen (1923 and 1925) devised three tests intended to supplement the Seashore battery. In the test of relative pitch, the subject has to compare 100 paired intervals and say whether the second is larger or smaller than the first. For rhythm he has to state whether or not two rhythmic patterns played on one note are the same, and, if different, whether the first or second phrase has been changed. In the tonal sequence test, the listener has to judge the relative merits of four possible endings to a melody. The evidence of validity provided by Schoen is based on a comparison of scores with teachers' estimates for only 10 pupils.

In England, Henry Lowery (1926; 1929) produced three tests. In one, two cadences are played and the listener has to judge whether the second is 'more or less complete than the first'. Cadence tests are difficult to apply to subjects without musical training owing to the difficulty of describing them and because, in any case, two chord cadences present a certain ambiguity of key. Lowery also worked out a tone memory test which required the subject to recognise a theme after certain changes, e.g. after transposition to another key, and a phrasing test which in fact involved memory to rather a high degree. The retest reliabilities obtained with 130 girls, aged 12–14, were quite promising (·75 and ·71). But Lowery's energies were later devoted to the work of directing a Technical College, where he succeeded in fostering active participation in music among his students. His tests were not further developed.

Also in England, James Mainwaring (1931) constructed tests of perception of pitch differences and rhythmic patterns, and of recall. His primary purpose was to study the cognitive processes involved in musical ability. He began with a consideration of the four physical attributes of sound. He assumed that everyone who could hear at all could perceive differences in loudness and distinguish for example between a saxophone and a harp. He therefore confined his attention to pitch and rhythm. For details of the tests, see Appendix I. Mainwaring, like Lowery, while remaining interested in the problems of assessing musical ability and in music education, never fully standardised his tests.

A battery of seven tests including four memory tests were developed by Otto Ortmann for use at the Peabody Conservatory of Music, Baltimore, but these have not been published.

Also in America Thurber Madison (1942) stressed the importance of the interval as the basic perceptual unit in music and carried out an extensive study of ability to discriminate intervals. As can be seen from Appendix I, his test correlated significantly with success in musical activities. Promising results were also obtained at Indiana University with his test of tonal imagery (Christy, 1956).

STANDARDISED TESTS

The Seashore Measures of Musical Talents

The Seashore Measures were the first standardised tests of musical ability to be published. Twenty years of intensive experimental work preceded the publication of the first edition in 1919.

Among the purposes of the Measures which Seashore listed in his book published in 1938 were:

1. to measure native and basic capacities in musical talent before training has begun, and, therefore, to make them independent of musical training;
2. to measure one specific capacity at a time; and
3. to make the procedure available for group measurement.

As can be seen from Appendix I the tests were of sensory capacities, rather than of musical abilities.

In interpreting the scores Seashore insisted that they should be used to provide a profile for each subtest and not totalled to give a composite score. In most cases, however, where the measures have been successfully used, for example at the Eastman School of Music in America and in the Rochester (NY) Public Schools, a general classification based on composite scores has been employed. Wing (Buros, 1959) recommends the use of the total score, if only because the reliability of the single tests is so much lower.

Seashore claimed that a reliability coefficient of over ·90 could be obtained on retesting under ideal laboratory conditions. In practice, however, much lower coefficients have been reported by users of the tests, at least for the 1919 version. The tonal memory test having up to five possible answers is the best from the point of view of reliability. McLeish (1950) added a third choice to the other tests by instructing his subjects not to guess when doubtful, but to record 'E' for equal. With this method he raised the reliability of the total score to ·90.

In the 1939 version the number of items in the subtests was reduced. Theoretically this should have reduced the reliability but results from the longer version were liable to be affected by inattention or fatigue, as Franklin (1956) has shown in a detailed analysis of some test and retest answers. Moreover, the more difficult items, where many subjects had to guess, have been eliminated. It is also now much easier for the subject not to lose his place, as the beginning of each column on the answer sheet is announced on the record.

Where the coefficients are relatively low, Seashore and his collaborators emphasise the importance of interpreting scores in broad categories only, and of retesting, if important decisions are to be based on doubtful performances. Retesting, however, may create other problems, in particular the later results may have been influenced by practice. In the case of musical subjects lower scores may be obtained, since they are liable to become bored.

While at least the measures of pitch, intensity and tonal memory can be regarded as reasonably reliable, the repeatedly expressed doubts as to the validity of the battery have still to be answered. The validity Seashore claimed for his tests was 'an internal validation in terms of success in the isolation of the factor measured and the degree of control of all other factors in the measurement' (Seashore, 1937). His critics, such as Mursell and Wing, do not deny that the pitch discrimination test, for example, is an objective and valid measure of sensory capacity. But they do question whether the results of such testing have much relevance to functional musical ability.

When the Eastman School of Music was opened, Seashore's first assistant, Dr Hazel Stanton, was appointed psychologist to the School, with full facilities for introducing a programme designed to validate the tests. After some experimentation, a method of classifying entrants into five classes: 'discouraged, doubtful, possible, probable, and safe' was worked out. Of the discouraged group, only 17% completed the four years' course in the standard time, compared with the 60% of the safe group who successfully graduated. Unfortunately, the predictive value of the Seashore tests alone cannot be determined from Stanton's reports, since an intelligence test was used as well as the Measures in the classification of entrants. No correlations are given, nor indications of the weightings attributed to each. The music tests seemed to have been considered the more important, since candidates with 'D' or 'E' grading were not admitted, whereas candidates with 'D' or 'E' gradings on intelligence were regarded as possible if their music gradings were high. To be classified as 'Safe' or 'Probable' the music talent grading had to be 'A' or 'B' (Stanton, 1935). This procedure, using the Seashore Measures interpreted in broad grades along with an intelligence test, has greatly improved the quality of the Eastman students.

Ruth Larson (1955) claimed similar results from a selection programme based on the Seashore tests which she had carried out for 25 years among the schoolchildren of Rochester, NY. Less favourable conclusions were reached, however, in an extensive study carried out at the Cincinnati College of Music, Ohio. There Taylor (1941) investigated how effectively the subtests of the Seashore and Kwalwasser–Dykema batteries (see p. 30) could forecast either success in a college of music or success in music as a profession. Marks for dictation, sight-singing, harmony and performance were used as criteria of college success. Professional success five years after leaving college was assessed by very careful enquiries from at least one person competent to judge. Where the information that could be obtained was inadequate, the case was excluded from the results. Compared with marks for dictation, all the Seashore test correlations were below ·30 and with sight-singing only Intensity reached the ·33 level with around 150 students. From the 93 cases graded into five groups from the highly successful to the complete failures as professional musicians, correlations of between ·34 and ·47 were obtained.

The validation studies of both the 1919 and the revised version of the measures when compared with empirical criteria, suggest that the tonal memory and pitch tests are the most satisfactory; but apart from a comparison with sight-singing scores by Salisbury and Smith (1929) which gave a coefficient of ·60 with pitch and ·65 with tonal memory, most of the correlations fall below ·50 (see Appendix I). Seashore himself protested against attempts to validate his measures against such criteria. It did indeed seem illogical to John McLeish (1950) to try to demonstrate the worth of a test that was supposed to provide a more reliable assessment of ability than teachers' opinions by comparing the scores with music grades and teachers' ratings. He therefore undertook a factorial study in which he tested a hundred students with the 1919 version and with the Wing and the Oregon tests (see below). He came to the conclusion that the measures were 'adequate for their original purpose, to measure the most elementary abilities required for the understanding and appreciation of music'. Comparing the Seashore and Wing batteries he concluded 'that Wing's tests measure much the same kind of ability as Seashore's but measure it at a higher or at least a different level, namely, that of musical meaning'. The measures, McLeish added, will be 'most effective if the scores are weighted in accordance with the calculated regression coefficients and if used in conjunction with other tests of musical appreciation'. He emphasised, however, the need for further validation studies (Buros, 1953, p. 343).

Kwalwasser–Dykema Music Tests

Jacob Kwalwasser and Peter Dykema published in 1930 a set of tests in which musical notes are used on the same lines as Seashore uses sensory material. Like Seashore's, the K–D battery contains measures of pitch, intensity, time, rhythm, timbre and tonal memory. Except for tonal memory, however, the corresponding tests in the two batteries do not appear to measure the same variables. To these, four tests have been added: tonal movement, melodic taste, pitch imagery and rhythmic imagery. The test manual does not mention reliability or validity. Studies on the reliability of the tests suggest that being shorter they are much less reliable than the Seashore tests. The most satisfactory seems to be the tonal movement test, the next best being tonal memory. But the reliability of the other subtests and even of the test as a whole are very low. We may wonder why the tests seem to be quite so unreliable, more so than other shorter tests. (For example see Arnold Bentley's short tests for younger children, p. 38 below.) No doubt many of the studies were carried out on select groups of music students, yet even with a wide range of musical levels, Sylvia Bienstock (1942) reported low reliability coefficients.

Because the battery is more musical and less tedious to take than the Seashore, it has enjoyed considerable popularity in the United States. Jack Holmes (1954) therefore thought it worthwhile to develop new directions together with a new set of weighted scoring keys and new norms. For example, where in the original instructions, the subject was asked to record S for 'same' or D for 'different', Holmes required the listeners to write E for 'equal' if the second playing was the same. If the subject noticed a difference, he was asked to write, for example, in the pitch test DH if he thought the altered note was higher, DL if lower. Holmes gave extra credit for a correct judgment of DH or DL. As a result of his revisions, Holmes obtained considerably improved coefficients with high school pupils.

While Holmes's procedure has improved the reliability of the tests their validity is still open to doubt. Lundin's (1967) table of validity studies reported by five different investigations shows 17 instances where a negative correlation was found between a subtest and the criterion and only four examples of validities of ·40 or over and one of ·59. For example the results of a study by Bienstock with over 100 students enrolled in the High School of Music & Art in New York showed that intelligence tests predicted individual success in the music courses better than did the K–D tests. This may have been because the tests do not contain enough discriminating items at certain important levels. Taylor found the pitch imagery and the K–D tonal memory tests the two outstanding subtests from the K–D and Seashore batteries in her investigation described above.

In his book Kwalwasser (1955) refers to several researches carried out by his students which show that his tests separate the most from the least musical children in a class, and that music students make appreciably higher scores than liberal arts students.

Farnsworth (1958) sums up the differing results of validity studies of the K–D tests in the words: 'Perhaps the modal forecast value for the battery as a whole would lie in the neighbourhood of ·40, with that for the individual tests being considerably lower.'

The Kwalwasser Music Talent Tests

In spite of the criticism of his earlier battery Kwalwasser brought out in 1953 a test that was supposed to measure thresholds for pitch, time, rhythm and loudness in only 10 minutes.

No reliability nor validity information was given in the test manual nor even standard directions for administering the test. From the few independent studies summarised in Appendix I the test in its present form appears to be too short to be reliable and too easy to be discriminating among older and more musical children.

The Drake Music Tests

Raleigh Drake, a musician as well as a psychologist, produced the first test based on musical material that had really satisfactory reliability and validity.

While reading for his Ph.D. at the University of London, he experimented with four tests: interval discrimination, retentivity, intuition and musical memory. In the retentivity test the subject is required to remember a musical interval, a beat given by a metronome, and a three-note sequence. He then has to judge whether each of several intervals is greater or smaller than the original one, whether a metronome beat is faster or slower than the original and whether a single note was the first, second or third note of the three-note sequence. (This ingenious test, intended 'as a test of absolute pitch or memory for isolated tones' seems to offer scope for the development of some really difficult musical puzzles!) The intuition test was supposed to measure 'intuitions' for phrase balance, time balance or key centre. When Drake experimented with these tests with four musical groups and one largely unmusical school group, he found that only the musical memory and interval discrimination tests gave satisfactory results with more than one group (Drake, 1933).

As these preliminary findings suggested that the musical memory test was the most promising, Drake, on his return to the United States of America, concentrated on standardising that test and in 1942 a recording was published. In 1954 Drake produced a rhythm test, which is in fact a test of whether or not the subject can keep a steady beat in his mind during a period of silence. This is of course an important ability for all types of musical performance.

The reliability of the two tests is high, especially for musical groups. (However Edwin Gordon reported some disturbingly large discrepancies between two testings of 20 subjects – see further, p. 161.) The advantage of measuring only two kinds of performance is that the subtests can be longer, thus improving reliability. The memory test alone takes 20–5 minutes, as compared with the 12 minutes required for the first three Wing tests. On the other hand, Drake's battery lacks a specific pitch discrimination test, though he was formerly of the opinion that a pitch test should always be included in any music testing programme.

The validity data given in the test manual is on the whole good, though the range of the coefficients is wide. Drake offers no explanation of the variations, apart from referring to the inaccuracies of the raters. Correcting for unreliability of the ratings would raise, for example, coefficients of ·70 to ·90. The validity figures obtained by Lundin (1949) and by Christy (1959) were much lower. John Ferrell (1961) of the

University of Iowa concluded from his study of the memory test with 180 pupils from three high schools that it successfully identified students that had superior musical aptitude.

The Drake memory test has survived for more than 30 years, and won considerable approval. The rhythm test is the only standardised test available that specifically measures ability to keep in time.

The Oregon Music Discrimination Tests

In 1930, Kate Hevner experimented with a test based on material from the compositions of accepted composers. The subjects had to listen to four versions of each item and judge which was the original, and which had been distorted by a mutilation of the rhythm, harmony or melody.

Keeping in mind four versions of a melody proved, however, too difficult a task for general use. In 1935, therefore, Hevner, in association with the Carnegie Foundation for the Improvement of Teaching in the Arts, devised and published an easier and more useful form of the test, where only one distorted version had to be compared with the original. Besides stating which version he prefers, the subject also has to decide which element – rhythm, harmony or melody – has been altered. Hevner found that the earlier test had considerable discriminatory value in distinguishing between psychology students and advanced music students and that results with the later form were similar. A second version of the test also requires the hearer to state the degree of confidence he feels in his judgments.

The Oregon tests have usually been regarded as tests of taste and appreciation, as distinguished from ear acuity tests. However, ability to perceive the differences between the accepted and distorted version is obviously required. Moreover, building up a listening repertoire of good music with which to compare the versions must partly depend on general auditory efficiency. McLeish found quite moderately high correlations between the test and both the Seashore and Wing tests. Of these three batteries, the Oregon seemed to demand the highest degree of musical ability, particularly the score based on judgment of the nature of the change. A procedure of assessing 'appreciation' of musical compositions which Hevner has also worked out will be described in the next section.

The Oregon tests enjoyed wide use and considerable esteem (see Lundin, 1958, and Farnsworth, 1958) for a number of years, though reports on experimental studies involving these tests have been few in number. In the 1950s, the records ceased to be commercially available.

However, Dr Newell Long of Indiana University has been working on a revision of the tests under the guidance of Hevner.

Thirty-one items were selected for reconstruction. To these were added 75 excerpts from piano, organ, string quartet and woodwind literature, mutilations being composed for each. A tape recording of these items was submitted to a panel of musicians who rated the items for probable difficulty and deleted those of doubtful validity. Two experimental forms of the test were then assembled and tried out on schoolchildren and college students. Promising results were obtained (Long, 1965). After further work has been done on the standardisation of the new tests, we may look forward to the publication of a new edition of these highly esteemed measures.

The Wing Standardised Tests of Musical Intelligence

Herbert Wing first started to work in the field of music tests in 1933. After a thorough survey of such tests as were then available, he decided 'to compile a comprehensive series of new tests, to assess their relative merits, and . . . to select a short series of proved diagnostic value'. There were 21 tests in the pilot survey. These were revised and later increased to 25. In addition to tests of a cognitive type, Wing sought to include tests of appreciation – 'the fundamental quality that all musicians would desire to find in any person who claims to have an interest in the art' (Wing, 1941*a*, p. 70).

Thirteen of the early tests were selected for recording on discs. The results were sufficiently satisfactory to encourage Wing to develop an even shorter form. After various modifications the seven most suitable tests were again recorded and then standardised. Further revisions have since been carried out, any item that appeared at all doubtful being removed or modified. The first three tests deal with aural acuity and the last four with taste or preference. The reliability of the whole test and of the first three subtests certainly seem to be good. The reliability of the four appreciation tests is less well established.

Considerable efforts have been made by Wing to establish the validity of his test. For example, he investigated the relationship between his test results and ability to persevere with the playing of a musical instrument. 333 boys, aged 14 to 16, were divided into Above Average, Average and Below Average groups according to their test scores. Wing (1948) then found that 40% of those with below average, and 27% of those with average ability, who had started to learn an instrument, had let their playing lapse, while only 2% of those of the above average group had ceased to play. A similar study of 718 adults showed that 83% of the below average group, 30% of the average group and only 9% of the highest ability group had given up playing (Wing, 1954).

Independent studies have confirmed the validity of the test (see Appendix I). Newton's (1959) study was carried out at the Admiralty with a view to reducing failure during training among the junior musicians at the Royal Marines School of Music. In his report Newton recommended that the test should be incorporated in the selection procedure and that a score of 70 (out of 136) should be adopted as a discretionary minimum. While candidates with lower scores would not necessarily be excluded, more stringent regard would be paid to their educational standard and personal qualities.

The Wing tests were included by R. R. Bentley (1955) of the University of South California, in a critical study of recently published tests. He matched 110 instrument-playing music students of a Californian High School with 110 non-instrument-playing music students on a basis of sex, IQ, grade placement and socio-economic status. He tested both groups with the Kwalwasser Music Talent tests, with those of Wing, of Whistler and Thorpe and of Gaston, and with the Farnum Music Notation test (see below). Of all the tests included in his study the Wing tests were the most discriminating between the instrument-playing music group and the non-instrument-playing group (many of whom had had lessons), and correlated almost perfectly with the total score of all the music tests. Correlation with an index of interest in music was higher than that of any other of the musical aptitude tests. Bentley concluded that where a very critical analysis of individual capacities is desired for guidance purposes, the Wing battery is the best test to use. When only a short time is available for testing, the first three Wing tests are the most satisfactory.

Though the battery was intended to be used as a whole to provide a general assessment of musical ability, some evidence is available on the relative value of the subtests:

Test 1. Chord Analysis. This has proved to be a most effective test over a wide range of aptitude. The opinion of Stumpf and Revesz as to the value of chord analysis as a diagnostic test of talent would seem to be justified. Even among the 41 professional students tested by Wing at the Eastman School of Music, it successfully separated the good student from the very good one (see Shuter, 1964, p. 386).

Tests 2 and 3. Pitch Change and Memory. McLeish (1950) regarded these as being the best validated of the whole battery, in the sense of showing the highest 'saturation' with a general musical factor. This also proved to be the case with a group of students of average musical ability, studied by the writer. However, among the highly select Eastman students these two tests were too easy to be really discriminating. Both Whittington (1957) and Newton found tests 2 and 3 to be among the three most effective. The pitch test was particularly good at picking out the junior musicians who were below average. The most efficient of all with the RMSM group was the memory test. 24 (out of 27) of the Above Average boys made above average scores, while 17 out of 28 Below Average boys scored lower than the mean of the total group of 223. In Bentley's results, the pitch test was the most effective single measure of pitch discrimination used in his study, and the memory test the most effective measure of memory.

Tests 4 to 7. Appreciation of Rhythm, of Harmony, of Intensity and of Phrasing. Wing found that the majority of the items of these four tests were too difficult for most children of nine years and under. Their usefulness increases with age and with level of musical ability. With the RMSM junior musicians the harmony test was second only to the memory test in efficiency in discriminating the good from the average and the weak from the average of the total group. The Eastman students, however, found it rather too easy to be highly discriminating.

The last two tests are especially liable to be affected by fatigue and loss of concentration, particularly with less talented students. However, in Wing's own factorial study (1941), the phrasing test gave the highest loading on a general music factor and Whittington found it one of the three most satisfactory for discriminating between his musical and unmusical groups.

The Gaston Test of Musicality

Thayer Gaston aimed at providing a general assessment of the subject's musical ability and interest in music. The latest version of this test, issued in 1958, presented all the tonal items on one continuous record.

The test consists of 40 items, the first 18 of which are in the form of a questionnaire seeking to assess interest in music. This leaves only 22 actual tonal test items. Reliability of the test is good. However, as Bentley points out, the validity evidence put forward by Gaston shows that the association between teachers' ratings and the scores reached a significant level only in the case of older children and of the total group studied by Gaston. The seven melodic memory items were, of all the tests Bentley investigated, most discriminating in distinguishing the instrument- from the non-instrument-playing group. The other items proved too easy for his subjects.

The Whistler & Thorpe Musical Aptitude Test

A piano version of this test was published in 1950 but so far no recording is available. Like Gaston, Whistler and Thorpe stress the use of musical material rather than the laboratory devices used by Seashore. The tests measure three aspects of music – rhythm, pitch and melody.

The reliability coefficients seem rather lower than those obtained with the Wing and with the Gaston tests. The range of validity coefficients is wide. The correlation of the total score with estimates of talent for instrumental performance was only ·52. However, in his review of the test, Wing (Buros, 1966) concludes that it may be sufficiently valid and reliable for ordinary school purposes with students of a wide range of ability, using a broad classification on a five point scale. He thought, however, that the percentile ranks for the separate tests might be misleading in implying a degree of accuracy which is not inherent in the separate tests as they stand.

The Gordon Musical Aptitude Profile

Published in 1965, this battery is the most sophisticated attempt to test musical ability that has so far appeared. More than six years of extensive and systematic research preceded publication. The musical examples are all original tunes composed by Gordon himself and played by a professional violinist and cellist.

The battery is intended to be administered on three days, and consists of three parts: Tonal Imagery (Melody and Harmony), Rhythm Imagery (Tempo and Metre) and Musical Sensitivity (Phrasing, Balance and Style).

The test called 'Style' in fact requires a judgment on which of two tempi is the more appropriate. The Phrasing and Style tests are designed to assess interpretative ability. The Balance test is supposedly related to melodic and rhythmic creative ability, at least indirectly. The ability to judge between two endings is no doubt a minor prerequisite of creative ability, though hardly any guarantee that the student could write an original melody with a suitable ending.

As we should expect with these longer tests, reliability is good. Some of the validity studies have given very promising results. In addition to the studies outlined in Appendix I, a longitudinal study was undertaken in 1963. All enrolled pupils – about 250 in all – in randomly selected elementary school classes (grade 4 or 5) are being given instrumental instruction over a three-year period. Each student was tested before the training began. At the end of each of the three years, their musical progress will be evaluated on three criteria: (1) scores on a test designed to assess ability to identify the musical notation of melodic, rhythmic, and harmonic passages; (2) ratings of tape-recorded performances of short musical passages, some prepared with the teacher's help, some without help, and some sight-read; and (3) on the teacher's evaluation of each student's progress compared with the other students in the group. Most of the correlations obtained at the end of the first year were in the region of ·3 and ·4. The total test score correlated ·6 with the composite score of the three criteria.

Perhaps the main disadvantage of the Gordon battery is its length. Though we agree that if musical aptitude is worth assessing at all, it is worth taking time to use the best possible measure, it remains to be seen how many teachers and research workers will in practice use the full battery. In particular, research is needed to investigate whether it is superior to the Wing tests.

The Bentley Measures of Musical Abilities

Published in 1966 these four tests were primarily intended for younger children (age 7 or 8 to 12).

The pitch discrimination test returns to the use of smaller than semitone differences as Seashore had done. A pilot test of pitch discrimination based on a comparison of melodic intervals from a semitone up to a tenth proved too easy. A possible means of increasing the difficulty would have been to mask the pitch change by adding concurrent notes, as in Wing's test. Since, however, the harmonic aspect of music seemed to have little appeal to younger children (cf. p. 83) and because artistic performance on pitch-variable instruments seemed to require subtle deviations from exact intonation comparable to *rubato*, Bentley decided to introduce smaller than semitone differences. In the current version, the 20 items range from one semitone (26 cycles per second difference at A = 440 cps) to 3 cps. His experiments with differences as low as 1 cps suggested that 3 cps was the smallest useful pitch difference that need be included in a group test.

Bentley's chord analysis test is similar to Wing's, although it contains a higher proportion of two note chords. Tonal memory and rhythm memory are tested separately. This makes the requirements of both tests less confusing for younger children. Only with a seven-year-old group did some of the children fail to understand the instructions.

The reliability at least of the whole test is satisfactory and the validity data are promising and it is already enjoying considerable use.

TESTS NOT COMMERCIALLY AVAILABLE

As long ago as 1944, Robert Lundin reported on some preliminary results obtained with a set of new tests. Further experiments were later carried out after various revisions had been made (Lundin, 1949). Lundin's aim was to measure in an objective fashion those aspects of music commonly taught in music theory courses. Theory in the USA includes not only written work but aural exercises such as writing melodies and chords from dictation. Unlike most other authors of musical ability tests, he does not purport to measure innate aptitude. As we shall see in Chapter XVIII below, Lundin believes that musical talent is largely the result of previously acquired skills rather than inherited capacities.

The five tests, intervals, transposition, melodic and rhythmic sequences, and a type of chord analysis cover quite a comprehensive range of musical abilities. The results Lundin obtained from his own experiments (see Appendix I), were on the whole promising. The reliability of the total score was certainly very satisfactory. The rhythmic sequence test gave low correlations with teachers' ratings. Lundin believed this might be due to his not being able to obtain an adequate criterion against which the teachers could rate the test. However, we might expect any useful measure of rhythm ought to show a closer relationship to performance than a correlation of ·17. The tests of melodic sequences, melodic transposition and rhythmic sequence seem to have been rather too easy for the music students tested by Lundin.

As the tests have not yet been published, independent evidence of their worth is not available. The interval test was, however, used by Faulds for his study of pitch perception (see p. 301) with satisfactory results. Lundin is at present working on a revision (private communication 1965), so that an interesting and satisfactory battery of tests may ultimately be produced.

Since 1950 the only two tests to cover a comprehensive range of musical abilities that have been published are the Gordon musical aptitude profile (described above) and the Aliferis achievement tests to be discussed in the next chapter. Test authors have tended to explore more specialised areas, such as the aesthetic and rhythmic aspects of musical ability.

The basis of Franklin's research (1956) at the University of Gothenburg was that a melody ends on the tonic. If the subject can find this tone, he has thereby demonstrated his musical ability. Franklin, therefore, sought to construct a series of short two-part melodies which would be interrupted immediately before the final tone, the subject then being required to complete the melody by singing the final note. After some experience of using this test of Tonal Musical Talent (TMT) in its individual form, Franklin constructed a group version. Though the music for the individual form of the tests has been published in Franklin's thesis no recorded version is yet available.

The reliability, in the 80s, seems very promising for a 15-minute test. The validity compared with a teachers' ranking was ·51. These coefficients refer to the individual form of the test. The group test is considered by Franklin himself as 'far from finished both with regard to reliability and validity', though usable to give some insight into the functioning of musical talent at a higher musical and psychological level. Faulds (1959), however, found the score of 35 unselected Princeton students averaged just over 1·5 points (out of 25) less than the mean score of 67 musical students from Westminster Choir. This may have been to some extent due to the sophisticated music students envisaging other acceptable endings and could perhaps have been avoided if the instructions had indicated that the required endings were in accordance with the idea that a tune should end on the tonic chord. However, the mean scores for both groups were considerably higher than for Franklin's Swedish students. The results of trying the test on younger groups are said to be 'quite promising' (Franklin – private communication). The test may thus prove to be more discriminating with subjects below College level.

Mueller (1956) commented on a need for assessing the intellectual processes involved in the appreciation of music and described a testing procedure. A complete composition is presented to the listeners and repeated three or four times. After the first presentation of the piece, the listener checks his answers to a list of questions. During two or three more hearings he continues to study the same list or more difficult lists. Such questions may be as simple or as difficult as the experimenter desires. Mueller describes an experiment in which over 100 students at Indiana University listened to the Third Movement of Mozart's *G minor symphony*. After listening to the piece the student was given five minutes to write a brief description of it. He was then asked to read through a list of 43 brief questions and to check with the number one his degree of assent on a scale as follows: strongly disagree, probably disagree, no opinion, probably agree, strongly agree. During the second and third playing of the Movement the student had the list in front of him and recorded his observations on each question with '2' and '3' to indicate second or third hearing, either during or after the hearing of the piece. The questions included 'piece includes three four time' and 'harp and piano are heard'.

The reliability of the scale was ·80; the correlation with music training was ·56 and with an interest in music scale ·70. Mueller's experiment produced interesting results and her procedure could usefully be adapted to obtaining information on the appreciation of many different types of composition with various sorts of listener.

George Kyme (1956) described a test of Aesthetic Judgment which requires the evaluation of paired performances, some taken from commercial recordings and others from recordings made at the Northern California Music Festival. The judgments were concerned with intonation and appropriate tone quality. For example, recordings lasting one minute each of soloists playing the same composition at the Music Festival were paired. In some instances one of the performances was simply duplicated. The subject, after hearing three matched performances indicated whether they were the same or different and if different, his preference. A judgment was required between the first and second and then between the first and third. When scored as a simple discrimination test of the detection of difference in the two performances, the correlation between the test and teachers' ratings was zero. But when scored as a test of aesthetic judgment its relation to the teachers' ratings ranged from ·56 to ·83 with an average of ·74. The test has not yet been published, but a copy for research purposes is available from the author.

James Hoffren (1964) of Jacksonville University has tried to produce a test of expressive performance of music that would resemble as closely as possible the judgment required in an actual musical situation. The ingredients of expression which Hoffren included were: rubato, smoothness, articulation, phrasing, unity, continuity, dynamic and agogic accentuation. Each test item consists of two versions of the same musical excerpt. One version of each pair is deficient in one or more of these elements of expression. The testee is asked to select the more appropriate version and he is not told which element of expression is lacking. Reliability coefficients range from ·53 to ·66. Hoffren's primary means of validation depended on agreement among judges drawn from the staff and graduate students of the University of Illinois School of Music as to which was the better version. The test correlated ·35 with the corresponding subtests of the Wing battery. Musicians as opposed to subjects with less training and experience made significantly superior than average scores. Hoffren is continuing to develop his test further.

Rupert Thackray (1966) of the Bedford College of Physical Education is experimenting with a battery of tests that cover a wide range of rhythmic abilities. From his experience as a music teacher he found that pupils have many different kinds of rhythmic difficulties. He therefore decided to devise several different sorts of rhythmic tests (see Appendix I). Besides aural tests he has tried to produce corresponding tests that could be administered visually and tactually. In order to compare rhythmic perception with rhythmic performance, he has also constructed performance tests which contained similar material to his rhythmic perception tests (see Chapter III).

Thackray's own experiments have already produced some interesting results. These will be discussed in Chapter XIX. The musical situations for which his tests have validity have yet to be established. If a standardised version is eventually published, they may prove useful in physical education and dancing, as well as in music.

III

Attainment, Performance and Interest

ATTAINMENT

Since measures of achievement have to be closely linked to what the students have been learning, teachers have often preferred to construct their own tests. But some standardised measures have been devised; in certain cases the authors have recommended that schools and colleges should establish their own norms rather than rely on those provided.

Some of the tests of attainment to be discussed in this chapter are so classified only because they require some previous training in music. Such tests as the Aliferis Music Achievement Test and the Farnum Music Notation Test require considerable aural aptitude, as well as a knowledge of notation. They might more properly have been included in the previous chapter. A number of pencil and paper tests of musical knowledge also include tonal items making classification difficult.

The best standardised test so far to be published is the *James Aliferis Music Achievement Test College Entrance Level* which appeared in 1954. Aliferis's aim is to assess the student's power of auditory visual discrimination, i.e. his ability to visualise the musical notation of what he hears, and to hear inwardly what he sees. The tests are divided into three sections: melody, rhythm and harmony. In each section there are both elements and idioms. By a melodic element, Aliferis means an interval. By a rhythmic element, he means a figure of one beat duration. By a melodic idiom, a four-note figure pattern. The rhythmic idioms consist of a combination of two rhythmic elements. When taking the test, the subject has to select, for example, which of four intervals is the one that is being played on the piano.

The standardisation of the test has been very thorough. Norms have been collected from different types of college in four regions of the United States. The user is intended to score the test according to whatever set of norms best matches his testees. It may be rather difficult for the non-American to decide which to use, but the differences are not very great. As we can see from the figures in Appendix I, the test appears to be of good reliability, except perhaps for the rhythmic section considered separately. The correlations with success at college music are satisfactory, although Aliferis exaggerates when he calls them 'high'.

Although the battery is called an achievement test it certainly measures much more than a knowledge of music notation. Wing indeed considers that it might well prove to be a sound diagnostic test of general musical aptitude at College Entrance level (see Buros, 1959). Aliferis himself suggests if it is to be used for predictive purposes, it should be supplemented by an audition, an intelligence test and the Seashore measures. But as it is much less taxing on auditory memory span than, for example, Drake's memory test or the Wing memory and appreciation tests, it might be more useful to supplement it with either of these rather than Seashore.

Encouraged by the success of his College Entrance tests Aliferis published in 1962 a test on similar lines for use at the end of the second college year. It includes comparison with the notation of harmonic elements (chords) and of melodic and rhythmic idioms. Norms are again provided for various types of colleges and regions in America. The reliability is quite good except again for the rhythm test considered on its own. The validity figures which Aliferis quotes in his manual are rather lower than those for the College Entrance level. Wing (Buros, 1966) in fact found its discriminatory power disappointing and considered that further research should be undertaken with a view to including more easy and more difficult items.

Stephen E. Farnum tried out several tests in which notation was compared with musical excerpts during his doctoral studies at Harvard. He developed a pilot form of such a test with 80 items. Experiments with 300 children aged about 13 showed that 29 items were satisfactory. After noting the musical problems involved in these satisfactory items, he devised 51 new items on similar lines. After further experiments the 40 most discriminating items were selected to form the *Farnum Music Notation Test* which was published in 1953. Each of the 40 melodic phrases is four bars long. One bar of each melody as played is different from the accompanying melody in notation. The subject has to mark the number of the bar in which the change has been made. The difference may be in pitch, in rhythm or in both. In fact, more than 75% of the changes are in pitch. The items cover a wide range of ability. Separate norms are given for boys and girls, and separate norms for pupils who have had music lessons. Farnum, however, suggests that users might profitably compile norms based on their own scholastic standard. As can be seen from Appendix I, the reliability and validity figures are highly satisfactory.

Among several tests mainly concerned with testing a knowledge of rudiments are the *Kwalwasser–Ruch Test of Musical Accomplishment* (1924) and the *Diagnostic Tests of Achievement in Music* by M. Lila Kotick and T. L. Torgerson (1950). The Kwalwasser–Ruch Test consists of ten parts, which were based on the recommendations of the Music Supervisors' National Council on curriculum in American schools. Test 3 requires detection of errors in pitch, and Test 10 the recognition of familiar tunes from the written notation. The rest of the items require purely factual knowledge of musical symbols and terms. The test was designed for use in Grades 4 to 12. Each of the ten parts has a definite time limit, the whole test requiring 40 minutes to complete. The norms are based on results obtained with over 5,000 pupils. A reliability of ·97 is claimed by the authors for the total score.

The ten parts of the Kotick–Torgerson Test cover a good sample of the various abilities that can be expected in the rudiments of music. It is intended for Grades 4 to 12. No norms are provided, since the authors recommend the use of local norms owing to the great differences in the standards of instruction found in different institutions. Wing (see Buros, 1960) points out that the instructions are fairly long, and require a certain amount of teaching. They thus depend for their efficiency on the explanations given by the person administering the test. However, Wing believes that the test should prove helpful to music teachers once they have gained experience of evaluating the results.

PERFORMANCE

The earliest attempt to assess performance in an objective manner was the *Hillbrand Sight-Singing Test* (1923). The individual being tested was required to sing six songs without accompaniment after studying them for a few moments. The performance was then scored on the basis of intonation, notes added or omitted, errors in time, etc. The range of difficulty was quite small. Two years later, the *Mosher Test of Individual Singing* appeared. Twelve items were presented to the subject, who then had to sing them back. The score was based on the number of bars sung correctly.

The *Watkins–Farnum Performance Scale* (1954) is the most important attempt to provide an objective grading of instrumental performance that has so far been published. It consists of 14 sight-reading exercises which are graded in difficulty. The easiest is intended for pupils who have only been studying the instrument for three months, while the most difficult would be an exacting test after several years of study. The system of scoring takes account of pitch and rhythm accuracy, correct tempo, the observation of expression marks, slurs and repeat signs. The reliability claimed is around ·9. The correlation between test scores and the ranking of students by their teachers had a median value of ·83. The instruments for which the scale is available include most of the woodwind and brass instruments, and the snare drum. It would be useful to have similar scales for the piano and for string instruments.

Robert Seashore (1926) devised a 'Rhythm Meter', a gramophone with contacts embedded in the turntable at various points. A number of different rhythms can be provided for the subject, who must try to make his taps on a telegraph key coincide with the clicks he is hearing. Nielson (1930) found significant correlations between this type of rhythmic performance and the rankings of superior compared with poor music students. Harold Williams (1933) adapted the device in order to study the motor rhythmic performance of young children.

Mira Stamback (1960) described a set of three rhythmic tests for small children. In the first test the child is asked to tap with a pencil on the table. Twenty-one taps are counted and timed with a stop watch. The time is recorded along with a note of any changes in speed or irregularities. In the second test the experimenter taps a pattern, his hand being screened from the child's view. The child then tries to reproduce the pattern. In the third test the child is shown written symbols of the pattern and asked 'How do you think this should be tapped?' The experimenter notes whether the symbols for long and short are understood or if both are tapped in the same way. Norms for the ages 6–12 are given.

As mentioned in Chapter II Thackray has developed a battery of tests concerned with rhythmic performance. The apparatus he used consisted of an improvised tambour (a piece of stout rubber stretched over the top of a tin can) attached by rubber tubing to a recording tambour fitted with a pen, which records on a kymograph. The tests are administered individually and as can be seen from Appendix I contain similar material to the tests of rhythmic perception.

INTEREST

As mentioned in the last chapter, part of Gaston's test of musicality is concerned with assessing the child's interest in music. The answer sheet for the Wing test also includes a number of questions on interest in music and previous experience of performing or singing. Other questionnaires have been developed, for example, by Rainbow (see p. 193), as a means of assessing interest in music for research purposes. In many cases such questionnaires will provide sufficient information. However, two or three scales have been devised specifically as objective measures of the subject's attitude to music and of his interest in music compared with his interest in other vocational or leisure time activities.

Kate Hevner and Robert Seashore (see Mueller *et al.*, 1934) adopted the method developed by Thurstone for the construction of aptitude scales. Their test is composed of 50 statements about music with which the subject is asked to agree or disagree. Examples of the statements are 'Living would be a much more dull and drab affair were it not for the beauties of music' and 'I believe the world would be just as well off if there were no music in it'. Each item on a Thurstone scale has been prejudged by a large group of people and rated as indicating a completely favourable, or a completely unfavourable attitude, or one that falls between these values. The reliability of the Seashore Hevner scale for college students is ·90. Farnsworth (1964), having found the scale valuable for research purposes, provided a new set of weights. Fifteen items received significantly different weights from the original.

Strong's Vocational Inventory Blank was developed at the Carnegie Institute of Technology. The items deal with the respondent's like or dislike for a wide variety of specific everyday activities, or types of person. It has been empirically keyed for different occupations on the assumption that there are differences of interest among persons engaged in different occupations. The relative frequency of a given response among, for example, engineers, as opposed to men-in-general, determines the weight given to each response, which can vary from minus 4 to plus 4. Considerable correlation has been found between the rating obtained with the blank and eventual choice of occupation. Reliabilities of ·8 or above have been found over a period of a few years, and of ·69 over 18 years. A scale for men and for women music teachers has now been developed, based on testing 500 male teachers and 450 women teachers. A scale for men and for women orchestral performers is also available.

The purpose of the *Kuder Preference Record* (1951) is to assess relative interest in a small number of broad areas rather than in specific occupations. In each item, the respondent marks which one of three activities he would most like, and which one he would least like. For example, he has to choose between the following three activities: browse in a library, watch a rehearsal of a large orchestra or visit an aquarium. Extensive item analyses have been carried out with high school and adult groups. Music is among the ten interests for which the results are scaled. Separate sex norms have been published for high school, college and adult groups. The reliability is around ·9 and results have generally been found to be stable over the period of a year. Among younger groups, shifts in high and low interest areas are relatively frequent, as might be expected.

IV

Conclusions

Though many problems remain, it is safe to conclude from the two preceding chapters that useful tests of aptitude and attainment have been developed. The most important contribution has been the evolution of prognostic tests.

While the test author can be expected to provide clear instructions for the administration of the test, norms based on an adequate sample, and empirical evidence of the test's reliability and validity, the user still has the responsibility of choosing the most suitable test available and of interpreting the results intelligently.

CHOICE OF MUSICAL ABILITY TESTS

The most suitable test for a particular purpose will partly depend on the age of the subjects and their probable musical level. For most purposes, especially for forecasting success with an instrument, the Wing tests appear still to be unrivalled. They cover a wide sample of useful abilities; their reliability and validity are high for tests of an aesthetic nature and, as will be seen in Chapters XV to XVII, they are relatively uninfluenced by previous training. If a shorter test is required, the first three Wing tests give very satisfactory results. In fact, Wing doubts whether it is worth giving the appreciation tests to younger children and suggests the use of the three tests as a first grading. For children younger than nine, Bentley's tests can be recommended. The Drake Memory test is another reputable test which is commercially available. If sufficient time is available, the Gordon Musical Aptitude Profile could be tried. The rhythm parts of the Gordon test, which take only 36 minutes might be used in conjunction with the first three Wing tests.

The above tests are based on musical material. Is there still a place for the Seashore Measures in the assessment of musical aptitude? McLeish, while agreeing that musical subjects would prefer the Wing tests, considered the Seashore battery would be more acceptable to the musically unsophisticated. In forming this opinion McLeish was no doubt influenced by his own experiment with University students (see p. 300). As students of psychology they may have found Seashore's tests of interest from the point of view of psychophysical methods. As his study was carried out shortly after the end of the war, a fair proportion probably consisted of older, ex-servicemen. For the purpose of discovering talent worth special training it would seem better in most cases to use tests likely to appeal to those with some liking for music. However, one may have to recognise that even musical children may have acquired a distaste for classical music. This is particularly liable to happen among boys, who too often regard music as an effeminate subject. The 'scientific' nature of the Seashore tests may make them more acceptable to boys. A liking for music has been known to grow in individuals who have discovered as the result of a test that they are gifted.

Since the tests seem to measure rather specific abilities, Lundin (1958) suggested we should 'find the specific performances where these abilities are most needed before we discard the Seashore tests as being useless measures of musical talent'. If we ask what type of musical activity is most closely related for example, to the pitch test, playing a stringed instrument with satisfactory intonation would seem an obvious area for investigation. Salisbury and Smith's results with sight-singing seemed promising; however, they were not confirmed by Taylor. Moreover, Seashore and Mount (1918) found little relationship between the pitch test and activities like singing a scale. Perhaps what we should be looking for is cut-off points rather than for linear relationships between performance on the Seashore measures and success in learning music.

The general level of the individual's musical ability is perhaps the first point to establish. Then, more specialised capacities might be probed further, if, for example, he is thinking of learning an instrument like the violin. The potential solo violinist will require a finer degree of pitch discrimination than the child whose general talent for music is not likely to take him beyond the second violins of the school orchestra. These are matters which require further research. Lundin also suggested the repetition of the sort of validation study carried out by Stanton at the Eastman School of Music. But other measures would be included as well as the Seashore and the results would be presented so that the specific contributions of the music tests, intelligence and case history, etc., could be judged.

CHOICE OF ATTAINMENT TESTS

The Aliferis Achievement Test College Entrance Level seems to give good results. For children, aged 11–13, the Farnum Music Notation Test is useful, if the published norms fit the user's requirements or if he can compile his own. For assessing performance on wind instruments, the Watkins–Farnum Performance scale is of value.

Standardised achievement tests are a useful means of evaluating the relative effectiveness of different programmes of musical training (see Colwell, 1963). A danger inherent in any achievement test is that the teacher may come to assume that an average score indicates not only what the pupil *has* achieved but what he *should* score.

INTERPRETATION OF RESULTS

Two points should be borne in mind:

1. Usually a high score is more trustworthy than a low one, for it means that the child has produced some positive evidence of musical aptitude. On the other hand, one of the main uses of aptitude tests is for screening out those who will not achieve success in music without enormous effort. However, low scores should be treated with caution, since some extraneous factor may have prevented the child from doing himself justice on the particular occasion.

2. Though the authors of most prognostic batteries have aimed at producing tests that are as little affected as possible by past experience of music, this aim may not have been fully attained. Lars-Gunner Holmstrom (1963) carried out an extensive study of the prognostic value of a simplified version of the three Wing aural acuity tests and a rhythm test of his own. He tested over 1,000 Swedish schoolchildren when they were eight to nine, and again two years later. He finally concluded that great problems arise in concrete prognostic situations from the differing effects of past experience on different tests. He considered that further research was needed into the effects of early music lessons on test results. This question will be examined in detail in Chapter XVII. It would certainly seem wise to interpret a child's score in the light of what is known about his past experience of music. Information on the musical status of his home may not be very exact. But, if an instrument has been studied, it is usually easy to find out how long the lessons continued. Some rough idea of the quality of the tuition may be judged from the pieces the child has learned and his general

technique – or even from the reputation of the teacher, if he practises locally.

Even if diagnostic tests do not provide the full answer to the question whether it is worth spending time and money on an individual's musical education, so long as they added some information of value they would be worthwhile. Though in Bienstock's research (see p. 31) neither the K–D tests nor an intelligence test were very effective in the prediction of individual success, the extent of previous training and an audition were even worse.

OTHER USES OF MUSICAL ABILITY TESTS

Farnsworth (1961) notes the usefulness of objective music tests to private teachers who, knowing full well that among their pupils were a number who would profit but little from further lessons, have made of the tests a dramatic way of proving to parents that money and time were both being wasted. In this instance the tests were not used as diagnostic tools but rather as objective proof of what was already painfully obvious to all but the doting parents. Wing (1948) points out that test results can help to prevent the teacher over-driving the pupil who is shown to be already working to his full capacity. They can also serve as a guide in identifying the child who is not reaching the standard indicated as possible by his test result, through a lack of diligence or some other reason.

Some of the tests primarily designed for the detection of musical aptitude have proved useful in other spheres. The Seashore pitch test is applicable wherever discrimination of fine differences is required, e.g. in some branches of the armed forces. The Seashore time test was shown in one study to offer 'prospects of being valuable both in the diagnosis of aphasics and the estimations of prognosis in aphasic patients' (Sievers, 1955). Stamback, too, found that children with reading difficulties and speech impediments tended to fail on her test of tapping back rhythms. Both a group of subnormal intelligence (mental ages ranged from 7:6 to 8:6) and a group of dyslexic children made scores inferior to a group of normal children of eight years old. Whereas the subnormals had more difficulty in Test Three (understanding the symbols), the dyslexic children had most difficulty with reproducing the patterns correctly.

PART II

The Development of Musical Ability

V

The Development of Musical Ability

'Music begins at birth. Babies have always been lulled to sleep by song or by their own crooning. Where there are many lullabies musical growth and interest are encouraged. In this day, when the crib is never far from the radio, the infant becomes as familiar with the sounds of music as with the sounds of the mother tongue. To shape these sounds into music is as easy as to shape the sounds of language into meaning' (M. Emett Wilson, 1951).

Since the tests described in Chapters II and III are not applicable to very young children, our knowledge of their responses to sounds and to music comes from case studies of individual infants and the more scientific observations and experiments with larger numbers of children.

Among the pioneers of child study were Wilhelm Preyer, Millicent Shinn and Wilhelm Stern, all of whom mention musical and auditory responses among their accounts of the first years of life of individual children. Although these writers' interpretations of their material may need modification in the light of subsequent research, their observations were carefully carried out and still have value. Preyer (1901) kept a complete diary of observations made at least three times a day, of his own son from birth till the end of the third year. Shinn's data were based on close observation of her niece, supplemented by material, also mostly biographical, from other sources. Stern and his wife recorded details of the development of their own three children through many years. All these children seem to have grown up in reasonably musical homes. Neither Preyer's son nor Shinn's niece seemed to show any marked talent in responding to its relatives' efforts to elicit musical responses. In later childhood Shinn's niece 'developed a fair average musical taste, and correct perception of time and pitch'. Unfortunately, this is the only child mentioned by Shinn for whom any evidence of later achievement in music is available.

Mary Shirley (1933) included some data on music in her study on the first two years of life, which was based on material from daily record sheets kept by the mothers and from frequent home visits to some 20 babies.

Of available case studies specifically concerned with music, Wing's account of the musical development of his two daughters is the most accurate and detailed. As 'opinions formed from an intensive study of these two children were checked by as much evidence as could be collected concerning other very young children', Wing considered that 'it may be reasonably assumed that the process of development is likely to follow along similar lines for different individuals' (1941*a*). The elder daughter was the more talented and has in fact become a graduate teacher of music, while the younger, though quite a competent cello player, has preferred to study medicine.

A disadvantage of biographical material is that the observer (however objective he tries to be), especially if he is also the parent, may tend to overestimate the child's abilities, or to interpret behaviour in the light of his own adult preconceptions. Such studies are also unrepresentative, since parents interested and intelligent enough to keep and publish records, tend to have superior ability.

In the case of music, at least one parent must have a sufficiently keen ear to be able to identify the child's responses. Apart from the possibility of talent being inherited by the children, such parents are likely to provide a musically stimulating environment. But, 'even with these limitations the intensive study of the single child has its advantages over the present more popular method of impersonal observation of masses of children' (Allport, 1937). One such advantage is that the child is observed in its normal home surroundings.

On the other hand, summaries of typical behaviour at various ages form valuable supplements to biographical material. The intensive work of Arnold Gesell and his collaborators in the United States has substantially added to our knowledge of average normal children, particularly during the first five years of life. Music was included among the cultural and creative activities which they list as being characteristic of children at various ages (Gesell and Ilg, 1946).

The reports of Gladys Moorhead and her associates at the Pillsbury Foundation School in California are an important source of data on the musical behaviour of young children. The School was set up to study 'the music of young children, to discover their natural forms of musical expression and to determine means of developing their musical capacities, particularly in the field of spontaneous creation' (Moorhead and Pond, 1941). Most of the children, whose ages ranged from one and a half to eight years, remained in the School for one to two years. Between

20 and 27 were on the register at any one time. The School was equipped with instruments chosen for simplicity, variety, intrinsic worth and adaptability to the purposes of the children. These included a number of Oriental instruments. In this carefully designed environment the children were left free to sing and play as they pleased with a minimum of guidance. As far as possible all the music produced by the children was noted, or recorded mechanically, and all activities which seemed to be musical were described. A very broad definition of music was adopted. The uses of the voice in speech were considered to be very similar to its use in song. All sounds produced by striking on hollow blocks or on the floor were regarded as sufficiently similar to the rhythmic patterns produced on percussion instruments to be counted as embryonically of musical value. The published reports do not say how the pupils were selected. Such a school would be likely to appeal to parents who were musical, or at least interested in music themselves. Three children whose individual progress is described in detail are stated to have come from good homes where they had the opportunity of hearing music; but none lived in an environment where music was a central interest. If these homes were typical, the children were likely to be rather above average in musical ability. It is not usually made clear in the reports at what ages the various activities occurred in the children.

THE WORK OF PIAGET

Current thinking in child development has been largely influenced by the ideas of Jean Piaget, the Swiss psychologist. Observations of his own children led Piaget to carry out many experiments particularly concerned with throwing light on the successive periods of development through which every child passes. Though he has been criticised for a lack of scientific rigour in his methods, many of his findings have been confirmed by other people repeating his experiments on children in other parts of the world.

The boundaries between the stages which Piaget postulates are not to be regarded as sharp dividing lines but as times when intellectual abilities are undergoing considerable change. The ages given are intended to be only approximate; what is more important is the sequence of stages.

The first period of life from birth to about two years of age is called by Piaget 'sensori-motor'. The first few weeks of this period are occupied by the modification and development of co-ordinations between the reflexes present at birth. More sustained behaviour directed to a goal follows later. Piaget emphasises not only the experiences of the child but also the child's ability to utilise these experiences in learning to adjust to and to deal with his environment.

The second 'pre-operational representation' stage includes two phases: the pre-conceptional phase from two to four and the intuitive phase from four to seven or eight. During this stage the child learns to form concepts of a reality more remote in time and space than immediate perceptions. At the pre-operational stage, judgments are intuitive, ego-centred, inconsistent and made in terms of single relationships. Perceptual dominance causes the child to focus on only one aspect of a situation to the exclusion of other aspects; hence he can deal with only one problem at a time and is unable to co-ordinate relationships. At the Pillsbury Foundation, for instance, Moorhead and Pond observed that 'to combine concepts of timbre and rhythm seems to be somewhat too complex for the small child'. When the child was excited about a rhythm he would pick up, say, the nearest drum whereas he would normally choose an instrument for its timbre with some care.

Operational thought emerges at about seven or eight when the basic stock of concepts becomes organised into coherent systems. An operation differs from simple action in that it is internalised and reversible. 'Internalised' means that the child can carry out trial and error in his mind without overt action. 'Reversible' means that the child can conceive a corresponding reverse procedure in regard to any set of activities, e.g. subtraction is seen to be the reverse of addition. But concrete thought remains attached to empirical reality. The child begins by acting, during the course of which he seeks to co-ordinate the sequence of results that he has obtained. Not till the stage of 'formal operations' is reached at 11 or 12, will he be able to reason deductively from first principles and test them empirically.

Central to Piaget's theory of concept development and essential for the appearance of the operational system of thought is the principle of conservation – the invariance of a given empirical factor throughout observed changes of state. Only when a given element remains permanent and independent of changes in its form can the mind use it in building a conceptual framework of the physical world. When the child can see for example that the total amount of a liquid remains the same even when poured into a glass of different shape, he 'conserves'. Piaget traced the development of conservation from a stage of non-conservation, where the child focuses on only one biasing aspect of the stimulus field, through an intermediate stage where the child oscillates back and forth between the major biasing aspect and a competing one, to exact conservation of the property concerned despite perceptual changes.

Piaget used visual or tactual-kinaesthetic tasks. But if conservation is a necessary condition for all rational activity, musical thought should be no exception. Marilyn Pflederer's research described on pp. 77–80, was therefore aimed at finding out how the principle of conservation might apply in the development of musical thought.

OTHER RESEARCH

Comparable to Piaget's belief in the dependence of later learning on the successful acquisition of earlier abilities is the interest that has been growing among child psychologists in the possibility of finding critical learning periods among human beings. From studies of animals there seem to be critical stages at which the organisms are ready for certain experiences. If suitable stimulation is not forthcoming from the environment, progress will be greatly impeded. Thus there is a brief and critical six-week period, at about the 11th month of life, during which the chaffinch develops its song pattern. If a growing chaffinch is reared out of hearing of all chaffinch song, the development of its song is greatly restricted. Once the critical learning period is over, the song is fixed for life, no matter how much the bird is exposed to the singing of other chaffinches afterwards (Thorpe, 1956).

Another method of trying to assess the importance of environmental stimulation at a very early age is by sensory deprivation experiments. For example, chimpanzees have been kept in complete darkness or without patterned visual stimulation for varying periods from birth, and their responses noted when they were brought out into the light (Riesen, 1950). The results suggest that there is in fact a critical age when the young animal is physiologically ready to respond to a class of stimuli either after one or two trials or a very short period of learning. If he is deprived of appropriate stimulation at this period, his development will proceed abnormally and learning to respond at a later stage will be difficult, or even impossible.

Such experiments cannot, of course, be carried out with human infants. But some evidence confirming the difficulty of making up for a lack of normal perceptual experiences early in life comes from reports from patients who have been born blind and later recovered their sight by operation. Von Sendon (1960) collected the available details of 66 such cases. Though criticisms have been directed at various aspects of this evidence, it does seem to be true that even the most intelligent and highly motivated of the patients needed a month to approximate to normal perception, even when their learning was confined to a small number of objects. As soon as they could see at all, they could perceive a single coherent object such as a square as distinct from a background and, perhaps, see a difference between two such figures shown together, but the differences were not easily remembered. Colours were learned comparatively quickly. No comparable reports exist about recovery from deafness; doubtless, similar difficulties would be encountered.

As far as musical development is concerned, there has been little attempt to formulate any critical periods. However, Desmond Sergeant (see p. 72), believes that exposure to music, particularly learning an instrument in the early years of childhood, is important for the development of absolute pitch.

VI

The Earliest Years

As early as one week after birth, the infant may stop feeding at the sound of a gong (Forbes and Forbes, 1927). In fact, even one month before birth babies have been known to move when a loud noise was sounded close to the mother. One woman is even said to have had to give up attending concerts because of the vigorous reactions of her unborn child to the music.

It used to be thought that newly born infants would respond to variations in the loudness and in the duration of sounds, but not to pitch differences. However, a recent study by Bridger (1961) suggests that the new-born infant can respond to tones differing in pitch. His subjects were 50 babies from one to five days. The tone of a given frequency was sounded till the babies ceased to show any motor response at all. Then a tone of equal loudness but of different frequency was sounded and produced in many babies an increase in movement and heart rate. One baby could discriminate between tones of 200 and 250 cps.

Peter Wolff (1963) of Boston, Mass., carried out a careful study in which he observed eight babies for four hours a day five days a week and for ten hours one day a week. He used a number of auditory stimuli in order to try to elicit smiling responses. During the first week of the infant's life he found that high-pitched voices were no more effective than the sound of a bell or a whistle at producing a smile. During the second week the high-pitched voice (which could be the mother's or that of a stranger or Wolff's own falsetto voice) began to produce more smiling responses than it did in the first week and more than other auditory stimuli. By the third week the infant would respond with a clear-cut, broad smile to a human voice if high pitched, more frequently than to a bell, whistle or rattle. Moreover, the infant would tend to go on responding for a longer period to high-pitched voices.

Mary Haller (1932) studied 19 children, aged three to five weeks. Using an audiometer, she attempted to classify their responses to several different pitches and intensities into those indicative of comfort and those interpreted as showing displeasure. Higher pitched and louder tones tended to evoke more restlessness than the softer, lower ones. One child whose parents were thought to be musical, showed great sensitivity to sounds of all kinds. Another who was quite unresponsive to the loudest tones of the audiometer displayed great interest in the footsteps of the nurse when hungry. But, even when the babies were awake, comfortable and newly fed, positive responses were given to only 47% of the stimuli.

Wolff also observed that while at birth the baby makes a variety of sounds, such as crying, grunting and squeaking, by four weeks gurgling and cooing sounds deep in the back of the throat are acquired. The early vocalisations of the child are spontaneous and do not reflect the speech around him. Thus the deaf child babbles for a period as if normal. Such babbling is important in establishing circuits in the infant's brain and nervous system, so that he learns certain movements of his vocal organs will produce certain sounds. In the case of speech, development to the next stage depends on the baby hearing speech sounds. The development from babbling all possible sounds to speaking one's native language is probably parallelled in music by the development from the earliest vocalisations to singing the notes of a specific musical scale.

Shinn's and Shirley's data on responses to music and other auditory stimuli are summarised below:

TABLE VI.I *Children's earliest responses to music and sounds*

	Shinn (*incl. Preyer*)			Shirley (p. 98)		
	n.	range in weeks	median in weeks	n.	range in weeks	median in weeks
Listens to voice				21	3– 8	4·0
Quietened by voice				15	2–11	5·0
Startled by sound				19	3–10	5·0
Pleasure at music	10	7– 7	6·3			
Listens to music				20	3–19	9·0
Coos or stops crying at music				18	12–32	15·0
Sounds purposefully made by hand	10	11–23	15·5			
Vocal sounds purposefully made	12	9–26	16·0			
Looks in direction of sound	15	11–21	16·0			

The considerable individual differences apparent in the above table might be largely due to the different standards applied by the various observers (in most cases the mother). The accounts of individual children show that the establishment of such an activity as looking in the direction of a sound involves a period of maturation and learning. Shinn observed it incipient in her nephew in the 13th week, but not unmistakable till the 19th. Though Preyer's son moved his head in the direction of the sound when his father knocked on a mirror behind him, no notice was taken of more distant sounds till the 16th week.

The table also shows that to the passive listening or diffuse responses of the nine-week-old baby more specialised and purposeful behaviour is gradually added. The earliest interest in sounds and music and attempts to create sound by manipulating objects with the hand or by vocalisations, are not, however, necessarily responses of a purely auditory nature. Unless Shinn's niece could see the noise-making process, sounds did not greatly interest her. When ringing a bell or rapping things together, she did not seem to care so much for the sound itself as for the relation of process and result. This suggests an interest embryonically scientific rather than musical. A similar comment might apply to Preyer's son who, at 10½ months, discovered by accident that the sound made by striking a plate with a spoon was dulled when he touched the plate with his other hand, and then proceeded to experiment with a spoon in the other hand.

As part of a research project at the State University of Florida, designed to compare the development of twins with singletons of the same age and sex, Gene Simons (1964) studied the responses to music of 12 pairs of twins and 12 pairs of singletons aged between 9 months and 2:7. The experiments were carried out in the infants' own homes. Responses to music were significantly less among the twins than any of the singletons. With piano recordings, responses of both were greatest to rhythmic music, less to melodic music, still less to harmonic music and least to dissonant music. Stimulating orchestral music elicited more responses than soothing orchestral music.

SPONTANEOUS MUSICAL ACTIVITIES

Shinn is probably right to think that a very young child's interest in sound-making activities is due to interest in producing effects by self-activity rather than in filling his ear with sound. Children's pleasure in making a noise may be largely due to the sense of command of their environment that they gain thereby. Yet, in addition, there appears to be some genuine interest in sound for many children. Moorhead and Pond claim that the small child 'produces sound from anything and everything around him. And while he does this he listens – it is not an aimless occupation. Some sounds please him more than others; some he will discontinue quickly, others he will repeat many times. . . . Music is, for young children, primarily the discovery of sound.' In the Pillsbury Foundation School it was, of course, much easier for the interest in sounds to acquire musical significance than in an ordinary home or nursery school.

The deepest interest of the Pillsbury children was in tone-colour. When the young child 'begins to use instruments for specific (e.g. dramatic) purposes he chooses the instrument whose timbre he considers most suitable'. The 29 children, aged between 6·5 and 7·5 years, whom Belaiew–Exemplarsky studied, still found their greatest joy in music in timbre or beautiful tone colour. Gesell and Ilg (1943) also noted that the 18-month-old child is 'very much aware of sounds such as bells, whistles, clocks', while the four-year-old 'likes to experiment with instruments, especially combinations of notes' on the piano.

Moorhead and Pond (1942) distinguish two types of music produced by young children, chant and song.

Chant appears to evolve from speech. In fact, the first type of chant is merely heightened speech; its rhythm is that of speech, but it differs from speech in that the most important syllable is strongly accented melodically. The second type of chant has a definite rhythmic pattern to which the words may be forced to conform.

The distinguishing characteristics of the second type are that it seems indifferent to melody; it is rigidly rhythmic and closely associated with physical movement; it is repeated through rises in intensity and pitch till a climax is reached, then stops. A chant may be started by an individual but is most often sung in groups. It occurs when the child is free and happy and it is immediate in emotional origin. The most frequently occurring occasion for chant (44 times out of 135) was motor activity of some sort. Four procedures in melodic construction which are 'absolutely fundamental to the child's musical concepts', as revealed by his chants, are quoted:

1. A falling minor third (probably slightly larger than a minor third of the tempered scale) after an accent;
2. An ascending, unaccented fourth preparatory to the accented upper note of the third a tone below it;
3. A leaning note resolved downwards; and (less frequent)

4. An ascent of a tone to an accent followed by a descending perfect fourth.

The following is an example of a chant in its 'final' form:

Arnold Bentley (1963) noted a rather similar chant, as sung by his own children. Gesell and Ilg mention the spontaneous singing of the minor third to such phrases as 'coal man, coal truck' as characteristic of the two-and-a-half-year-old child, while at four the child 'creates song during play – often teases others on a variation of the minor third'. At 18 months Shinn's niece began to amuse herself with a sort of tuneless chanting of syllables, which became a regular expression of happiness while she played. From a monotone during the second year, this later grew more varied and rhythmic till it had quite a pleasing and musical effect, contrasting significantly with her reluctant and ridiculous efforts at civilised songs. She apparently outgrew her own 'compositions' as she became more proficient at learning the songs which were sung to her. (Much of this spontaneous music-making was probably more in the nature of 'Song' than of 'Chant'.)

Moorhead and Pond considered that chant is 'the most primitive musical art form, for such it is *sui generis*, to be found among children and, indeed, among men in general. It is part of the living experience of primitive peoples everywhere, . . . as a primitive, pagan, unsophisticated musical expression *arising from those things which the child feels instinctively to demand such expression*.' They collected 150 chants over four years, but they do not say whether certain children were usually the initiators, and the rest merely imitators. It would seem advisable to treat attempts to draw parallels between the child's musical utterances and the music of exotic races with caution pending further investigation.

Besides chants, children will produce songs of their own invention. In contrast to chants, 'song is essentially produced by the child for himself'. These experiments in melody are changed and developed as the child wishes. The rhythm is free and flexible and if accompanied on a drum the song may be at a different tempo. The melodies are not necessarily diatonic and most seem not to relate to any observable tonal centre or not to a tonal centre found in Western music. They tend to progress by small steps, but large intervals are sometimes used to considerable dramatic purpose (Moorhead and Pond, 1942). In children with musical talent, the singing of self-invented melodies that exhibit the first beginnings of musical form can occur as early as the fourth year (Revesz, 1953). In these melodies, one finds uncertainty in general pitch, intervals, keys and motives, but some children reveal a sense of tonality. Stern's daughter at 3:8 would sing continuously for a long time. Her song was by no means tuneful, but sometimes there appeared a fragment of real melody, such as the beginning of 'Ring a Ring o' Roses'.

W. Platt (1933) wrote down some interesting examples of musical phrases that he had heard during 30 years of observing babies and toddlers. He had found some infants of only three months who could repeat at correct pitch the last note of a piece of music. At three months too, the infant might add a second note, usually a fourth lower, making a cadence from d^1 to s. One musical baby of seven months repeatedly expressed joy by crooning an octave. The upper note was uttered first with an indrawn breath and the lower note, quite in tune, with the outgoing breath. Platt noted a 'tune' on CGC^1 from a baby as young as four months, in imitation of a bark of a dog. He also quotes a little tune, lasting 22 bars, made up by a girl of 1:5, with the rumble of the car in which she was riding acting as a bass note.

The resemblance of songs produced by young children to existing music may be partly a function of the kind of music the child has heard, as well as his musical capacity. Thus, no attempt was made by Doreen Bland (1936) to make her son memorise rhymes or songs. When at four he started to sing his own songs, the words were about trees, sky and floods. The tunes bore no resemblance to recognisable airs, but preserved an obvious rhythm which was the only clue to any possible punctuation. He would sing intermittently when alone, or when walking with, but detached from, his parents. However, compositions of very young children are usually largely imitative, rather than creative. For example, at Pillsbury, Jay, a boy of 4:7 is reported to have 'devoted himself to a prolonged period of exact repetition', of a Mexican recording he had heard several times. 'He then attempted a more complicated pattern from the same recording, but was not able to establish it and returned to the first pattern for another long period' of exact repetition (see Moorhead, Sandvik and Wight, 1951, p. 21).

Moorhead and Pond believed that enforced conformity to the conventions of Western music before the child has sufficient background to see them in proper perspective is likely to hinder the growth of vital musical conceptual patterns.

They claim that eventually 'our children will tend to feel our system, which is continually being transformed and expanded, is closest to them . . . but their experience, knowledge and perceptive potentialities will be greater than if their experiences had been confined to that system' (Moorhead and Pond, 1942, p. 19). It is difficult to assess the educational value of the Pillsbury type of experience (as opposed to its psychological interest) in the absence of evidence showing whether children brought up in such a school do in fact end up with a wider appreciation of music than those who have been encouraged to listen to and sing a wide variety of 'conventional' melodies. Possibly what really matters is that the child should grow up in an environment where music is valued, and that he should hear music that is good of its kind.

DEVELOPMENT OF MELODIC SKILLS

The following table shows the ages quoted for the appearance of various melodic skills, by Wing, by Shinn and by Stern:

TABLE VI.2 *Appearance of various melodic skills in young children*

Recognising a tune	0:9	Wing	1st child
	1:0	Wing	2nd child
	1:8	Tilley	2nd child
	1:9	Tilley	1st child
	2:4	Shinn	niece
Reproducing a note correctly	0:9 (singing correctly every note from piano)	Friedemann (cited by Preyer)	
	1:1 (last note of song)	Wing	1st child
	1:8	Wing	2nd child
	2:4	Tilley	
	Not till 4th year	Shinn	niece
Repeating interval	0:8 (notes of cuckoo clock)	Safford	
	1:2	Wing	1st child
Some memory for general melodic shape	1:4 to 1:6	Wing	both children

Singing a tune correctly	1:3 (several songs)	Helliwell	
	1:4	Wing	1st child
	'Before they could talk'	Friedemann	3 siblings
	1:11 (one song, correct rhythm, fairly in tune)	Stern	son
	2:0	Wing	2nd child
	(nearly) 3:0	Preyer,	Daniels
	Not before 4th year	Shinn	niece
	5:6 (a few songs quite correctly)	Stern	son

It must be rather difficult to judge how far a very young child really recognises a tune. Wing states that his daughter appeared to recognise such tunes as 'Pat-a-Cake' by making the appropriate physical signs of clapping. That the response was to the melody rather than to the sound of the words was confirmed by the child being able to sing later on, before she could talk, a complete nursery rhyme with distorted words. By singing a song completely Wing meant with the steps and leaps in the right places though, for example, a fifth might be sung instead of a sixth. As evidence of the early development of some differentiation between tunes and some memory for their general shape, Wing observed that some children who did not sing (probably because their parents did not sing to them) quickly learned to indicate which of certain known gramophone records they wished to hear from the colour of the label.

It is interesting that it was the general melodic shape rather than the longest or most prominent notes that was learned first. This suggests that learning proceeds from an initial apprehension of a tune as a whole, with more definite perception of the parts taking place later. If knowledge of the tune was being built up by first learning its elements, one might have expected the most prominent notes to have been learned first (see further Chapter VII below).

The elder of the Wing children showed continual progress with age. Between two and three she could sing through the tunes of the *Children's Overture* with all its changes of key entirely without being prompted. She evidently had a considerable memory for melody, as she would indicate a record she wanted to hear by singing its theme and could sing the second phrase of a tune, if the first were sung to her. At 3:7 she sang through an Easter hymn which she had learned seven months previously and which she had not heard or sung, to the knowledge of her parents, since then.

Valentine (1956) contrasted infants under three who could sing accurately a considerable number of simple tunes with his own five children's inability to sing reliably a tune at seven or eight. Only one was able to do so even when adolescent; yet all his children heard a good deal of music at home, in church and in school.

Simons (1964) experimented to see how well his subjects could imitate short rhythmic patterns, melodic intervals, brief phrases and a song 'Teddy Bear' sung to 'da-da'. Each was given four times in succession and scored 0 (no response) to 4 (response identical or practically identical with the original). The twins aged nine months scored nil for the song, 2·00 for rhythm, 1·00 for pitch and 1·00 for the phrase. The singletons of the same age scored 1·00 for the song, 6·00 for rhythm, 10·3 for pitch and 5·85 for the phrase. Very considerable differences were found among the individual pairs, some of the younger ones making relatively high scores.

Gesell and Ilg (1943) record the following developmental stages:

2 years Sings phrases of song, generally not on pitch.
2½ years All of parts of several songs sung spontaneously at home or school.
3 years Whole songs reproduced though generally not on pitch; can recognise several melodies; is beginning to match simple tunes.
4 years A few can sing entire songs correctly.

Further evidence of the kind of melodic tasks which children between two and five can undertake comes from several studies of pre-school children, mostly concerned with how far such skills can be improved by training (see also p. 159 below).

Updegraff, Heileger and Learned (1938) obtained the following average scores from pre-school children, whose mean IQ was around 121, but who had had little training in singing. Four trials were allowed for each item of the tests.

Many of the five-year-old children were able to make perfect scores at singing one note and an interval. Out of 1,282 intervals sung by four-year-olds only 0·5 were in the wrong direction. The children's skill was not, however, too firmly established, as may be judged from the fact that they appeared unable to sing a note played on an instrument, though they were perfectly capable of reproducing it after it had been sung to them. Moreover, especially with the three-year-old group,

	Max. score	Aver. score at 3 n. = 16	Aver. score at 4 n. = 14	Aver. score at 5 n. = 36
Singing one note after it had been sung	9	6½	5	8
Singing an interval after it had been sung	12	7¾	6¾	10
Singing a phrase after it had been sung	14		10	
Singing a phrase after it had been sung	29			13¼

performance was inconsistent. On one day their scores might be high, while on the following day they seemed unable to sing so well.

Twenty-three pre-school children, median age 4:5 and with a mean IQ of 123, were trained by Drexler (1938) for 15 periods of a quarter of an hour to sing two songs previously unknown to them. 6 out of the 23 were then able to record an absolutely accurate version of the songs. The following table shows the percentages of boys and of girls who were able to learn to sing a scale at pre-school ages, as reported by Monroe (1903) from the investigations of his students, who may have applied slightly different standards of judgment:

Age	n.	Boys	Girls
2 to 3	38	29%	49%
3 to 4	64	31%	54%
4 to 5	46	34%	59%
5 to 6	12	40%	71%

Some data were also collected on the children's memory for the scale and for songs learned, after two weeks:

	Boys		Girls	
Ages	Scale	Songs	Scale	Songs
2 to 3	19%	43%	23%	59%
3 to 4	27%	50%	33%	61%
4 to 5	29%	47%	45%	62%
5 to 6	40%	50%	57%	71%

The songs may have been recalled better due to interest in the words, or because a tune is more musically interesting than a scale.

As we noted on page 68 very young children can often sing the general shape of a tune, even though the individual intervals are not exact. Teplov (1966, p. 199) quotes several Russian and German researches which confirm that there are two stages in the development of an ear for melody. At the first stage the child can recognise and reproduce only the melodic contour. According to Teplov perception of pitch has not yet been differentiated from perception of timbre. At the second stage the child can recognise and reproduce not only the direction of the movement of the notes, but also the correct intervals between the notes.

The first stage is attained by all children. If the second stage is reached easily, it is a sign that the child has a good ear for melody. Teplov emphasises that judgment of intervals grows out of a sense of tonality. Sense of tonality depends on perceiving that certain notes are 'stable', and give a feeling of completion when a tune ends on one of them.

In one investigation most children between the ages of 8 and 11 without any particular musical training were able to complete a tune on a stable note (third or fifth degree of the scale) if not on the tonic. Given some instruction in solfa, children of eight can perceive the difference between 'completed' and 'unfinished' melody, though they may not all be able to sing the key note to complete a tune. Franklin (1956), also, believed that an ear for tonality first begins to be established between six and nine years. Until an ear for melody and tonality is well established there can be no real appreciation of harmony.

DEVELOPMENT OF ABSOLUTE PITCH

According to Teplov (1966, p. 166) in a great majority of cases absolute pitch is shown very early, usually very shortly after the child has learnt the names of the notes. Thus absolute pitch has been noted in children of five or even four or three years old.

Among the cases Teplov quoted was that of a boy who had received his first piano lesson at the age of five from his father. When he had gone to bed, he would listen to his father playing the piano. One day when he was six, to his father's astonishment, he named certain notes which his father had played wrong. Another boy was able to name 92% of the sounds played to him on the piano when he was only seven years old.

Gebhart (cited Teplov, p. 169) made a special study of a very gifted boy. At the age of 1:9 this child could distinguish the sound of the piano from that of the violin. At 3:2 his mother played a C on the piano and told the child the name of that note. The next day he recognised it from among a series of other notes. By the age of three and a half he already knew the names of all the notes in one octave and could recognise them even after not having heard the piano played for several weeks. When tested by Gebhart at the age of five and a half he made only one mistake and that was due to the piano note having been badly tuned. During the tests the child recognised most of the notes without hesitation and rarely needed to sing them. He could also reproduce with great ease on the piano notes played on stringed instruments. When he was tested again at six, he named thirty-seven notes from a wide range of registers in 30 seconds. He could also recognise the pitch of other sounds he heard, like the buzzing of a bee or the note given out by his balloon.

Although the possession of absolute pitch is not a necessary component of a high degree of musical talent, it occurs much more frequently among professional musicians. Bachem (1940) pointed out that this seemed specially true of musicians with early music training. He goes on to say that attention to musical tones in youth seems to play a predominant role in the development of absolute pitch. Desmond Sergeant (1967) has recently carried out two surveys on the development of absolute pitch among musicians and music students. In his first survey he compared four groups of varying levels of musical talent and attainment. Group 1 consisted of 36 members of the professorial staff of the Royal College of Music, Group 2 of 30 experienced teachers from the Junior Department of the RCM. Group 3 was made up of 145 students in training at the RCM and Group 4 of 50 students training to be teachers who were studying music as part of their course but who would not have been able to reach the standard required for entry to the RCM. 69·4% of Group 1, 60% of Group 2 and 33% of Group 3 claimed to have absolute pitch. None of the Training College students claimed to possess absolute pitch. Sergeant believes that the age at which music lessons are begun is a critical factor in whether or not the child will develop absolute pitch. As evidence he quotes the following figures:

TABLE VI.3 *Age of commencement of music lessons*

Group no.	Absolute pitch subjects		Non-absolute pitch subjects	
	Mean	Mode	Mean	Mode
1	5·4 yrs	4 yrs	8·7 yrs	9–10 yrs
2	5·9 yrs	6 yrs	7·3 yrs	7 yrs
3	6·4 yrs	6·4 yrs	8·8 yrs	8 yrs
4	nil	nil	9·9 yrs	11 yrs
total:	5·9 yrs	5 yrs	8·7 yrs	7–8 yrs

These figures were supported by a second survey in which information was collected by questionnaires from over 1,500 members of the Incorporated Society of Musicians. The sample represented nearly 35% of the total membership and closely resembled the total membership in percentage of male and female subjects, musical occupation and the instruments they played.

Among members who had absolute pitch the commonest age for beginning musical training was 6 and the average age 6·1, the most usual age for beginning training among those who did not possess absolute pitch was 7, the average being 7·9. As can be seen from Fig. 1,

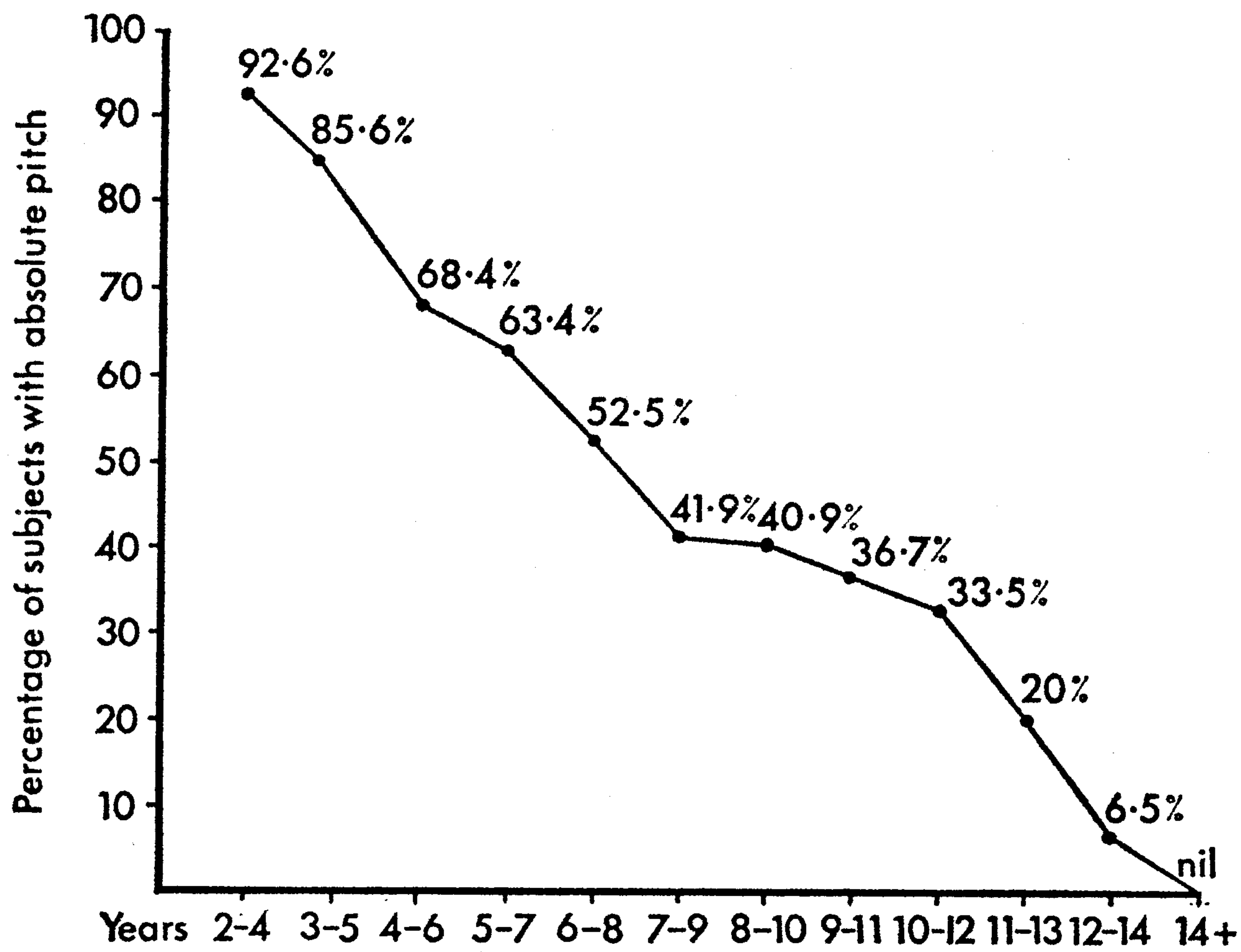


FIG. 1. Relation of age of commencement of training to absolute pitch.

the earlier training was commenced the higher the proportion of subjects with absolute pitch.

Sergeant checked the validity of his questionnaire data by actually testing a representative group of subjects. He devised a pitch naming test which consisted of a series of 50 notes. The subjects were not required to state in which octave the notes lay, only to give the letter name of each. Groups of five notes were played on each of 10 different instruments, each note being played twice. Results with the test given to 96 students at the Royal College of Music showed that:

'(*a*) The highest scores were always obtained on notes from instruments with which the subject was familiar.

(*b*) The highest score in all cases except one was obtained on the instrument which had chronologically been the first to be learned.

(*c*) There were many cases where the subject had changed from the instrument on which he commenced his musical training to some other instrument. In all these cases a higher score was obtained on notes from the instrument learned first than on the one which was now of greater importance in the musical life of the subject. In some of these, complete accuracy was shown on the instrument of childhood, and not a single correct judgement on the one later adopted.'

THE DEVELOPMENT OF RHYTHMIC SKILLS

There appears to be some difference of opinion as to whether rhythmic skills develop before or after melodic skills, or independently.

Revesz (1953) believed that in the period between the second and fourth years 'music and movement go together and cannot be divorced one from the other. In this period, for biological reasons, rhythm seems more important than melody.' Arnold Bentley (1966) also considers that as far as the child's ability to join in with a group is concerned singing in unison seems to occur less spontaneously and at a later stage, than coalescence upon a rhythm. Stern thought that rhythm makes a special impression on young children because, unlike melody, it can be grasped and copied by all sensory and motor organs.

On the other hand, Wing, while agreeing that the young child delights in physical activity, doubts its value from a purely musical point of view. In his opinion the first aspect of music to develop in the case of many children is melodic shape. When his daughter could reproduce a tune perfectly as regards melody, she still tended to shorten the longer notes considerably through a lack of patience and rhythmic sensitivity. At the age of three and a half or four, many songs still had notes shortened or lengthened without regard to the regularity of the beat.  would be sung, for example, instead of  From his observations of children engaged in music and movement in schools and in spontaneous play, Wing concluded that they frequently sing in tune for example 'Ring a Ring o' Roses', but do not normally skip or walk in time with the music.

It is in fact extremely difficult to judge whether children walking or marching to music are in fact keeping in time. Heinlein (1929) found that three experienced musicians playing from memory and watching children move 'in time' with the music tended to fit their timing to the movements of the children rather than vice versa. Heinlein then asked eight children aged three to five to walk 'in time to the music' on a runway which had electric contacts designed to record their steps. The beat of the music was recorded at the same time. Only one of the eight subjects showed any degree of co-ordination between the walking movements and the musical beat. This boy was the child who had appeared during the preliminary training period to be the least able to perform the task. Another boy who had seemed extremely responsive was much stimulated to motor activity by the music, but not once did his walking synchronise with the beat.

These results were confirmed by Jersild and Bienstock (1935) who used photographic methods of recording how exactly children under six could keep time in walking, or with hand movements. The following were the average scores out of a possible 200: at two, 41·8; at three, 56·3; at four, 82·3, and at five 97·5. Seventeen adults earned an average score of 174, some making perfect scores. Jersild and Bienstock found that their subjective ratings of the children's accuracy in keeping time were quite untrustworthy when compared with objective records.

Certainly some caution is needed when looking at the reports of ages at which various children are reported to have been able to keep time by movements. At one year Chapman's child (Shinn, 1907) could keep time and by 1:8 'could keep the polka step for six measures at a time'. One year is the age which Wells (Shinn, 1907), herself a musician, gave for her child being able to keep time (Shinn, p. 191). Towards the end of his second year Preyer's son could 'beat time with tolerable correctness' while trying to sing over a song. 1:10 and 2:8 are other times quoted. Shinn's niece could not be taught at any time during the first three years to keep time to music. According to Gesell and Ilg (1943), three-year-olds may be expected to 'gallop, jump, walk and run in fairly good time to music'.

Il'ina (1961) in Russia asked 130 children, aged 3 to 11, to move to music played in whatever way the music suggested. The older children were better able to correct themselves and change with the music. The younger ones tended to respond to changes in intensity with changes of tempo.

A much simpler task both from the child's point of view and that of the observer is producing a regular monotonous beat of his own. The first musical attempts of a child are characterised by a 'regular, unaccented beating', probably physical in origin, according to Moorhead and Pond (1942). Thus, Carl, aged 3:8, who attended the Pillsbury School for four months, 'reverted to the usual regular beat, perfectly even, fast and insistent'. Roy, aged 4:2, 'used a regularly spaced series of beats, with occasional pauses, moving arm and hand with pendulum-like regularity'. A little later the child will begin to introduce 'accentuation within the regular series of beats which he has set up, such accentuation being most often irregular. His rhythms, therefore, are not repetitive nor necessarily symmetrical, but their structure almost inevitably is related to the fundamental pulsation.'

95% of all spontaneous rhythmic groupings played by Simons' subjects consisted of notes of equal value. Among the other patterns that were produced, Simons quotes the following:  (by a girl singleton of 10 months) and  (by a girl twin of one year old).

Capacity to produce a regular beat or a rhythmic pattern on one's own is not, of course, inconsistent with finding difficulty in synchronising with the beat of others. Moorhead and Pond remark that 'the child has great ability to maintain his own rhythmic concepts against all competition and interference. When he plays simultaneously with other children each child is likely to go his own way. But in a long established, well-integrated group, the children's music is likely to assume the characteristics of a kind of rhythmic polyphony based upon the fundamental co-ordinating pulsation which they feel.' As among African drummers, simultaneous groupings of two against three are frequently found. The difficulty that the young child finds in conforming to a time pattern, or a beat, or a pitch outside himself can be considered part of the egocentricity of the very young.

To sum up, the first essential condition favouring the full development of musical ability in the earliest years is opportunity to hear music. No doubt it is highly desirable that some of this music should be made by a parent or other person of whom the child is fond. But if the parents cannot play or sing, they can at least try to listen to music with enjoyment and attention, and show appreciation of music and musicians. Before he is old enough for formal music lessons, the child can be encouraged to make his own music, with his own voice and with whatever sound-producing facilities his environment can offer, even if these fall far short of those provided by the Pillsbury Foundation. As Mursell and Glenn (1931) suggest, the child's interest in and love of tone can be the first growing point of his musical life. Parents will certainly agree with these writers in recommending that he should be encouraged to enjoy producing and listening to *beautiful* tone quality!

Possibly what is most important of all in these earliest years is that the child should come to think of music as a joyful activity.

VII

The Middle Years of Childhood

In the child over six as in the child under six, musical ability increases with age and, within each age group, marked differences of ability between individuals are found. Our knowledge of the development of musical ability is on rather surer foundations from the age of seven or eight onwards, since the results from the group application of objective tests become available. Before discussing the data obtained with tests, however, mention should be made of an interesting attempt by Pflederer (1964) of the Northwestern University to investigate children's responses to musical tasks embodying Piaget's principles of conservation. She devised musical tasks to study the conservation of metre, relative pitch and rhythm, at the age levels of five and of eight, and experimented with eight subjects at each age. To arouse the child's interest each task was introduced with a story. The responses to the task and to questioning were recorded on tape.

The first task was designed to study ability to conserve metre against the duration of notes. The children were simply asked to say whether six examples played on a drum and two tunes played on the piano were in 'two or three', i.e. duple or triple time. Unfortunately, the tune in triple time, *Lavender's Blue*, may have been familiar to the eight-year-olds. Only 29% of the responses given by the five-year-olds on the first hearing of drum examples were correct. Their verbal explanations were marked by an absence of conservation – their perception centred on the number of notes; there was no attempt to co-ordinate differences in durational values. 54% of the eight-year-olds' responses were correct. Five seemed to be at the intermediate stage of conservation, while three were approaching absolute conservation of metre. Both groups found the tunes easier – 44% of the responses by the five-year-olds, and 75% by the eight-year-olds, were correct. Three of the five-year-old children arrived at the correct answers through counting and clapping with the music. Two showed a definite absence of conservation; they tried to count out all the notes heard. The eight-year-old children showed evidence of conservation.

To study conservation of rhythmic pattern, four examples were played on a single-tone bell. The task was presented within a story context. The children were first asked to decide which one of four make-believe children played the pattern wrongly. The same patterns were then played on different tone bells to see if the children could conserve the rhythm patterns when they were presented in varying tonal contexts. The five-year-old group had difficulty conserving the rhythms under deformation of tone, since their perception concentrated on the tones. The eight-year-old children also found the second part of the task more difficult than identifying the pattern when played on the single-tone bell. Around 90% of their responses to the first part were correct, as against 75% for the second part.

Conservation of melody when the duration of the notes was changed was studied next. A four-bar phrase was immediately followed by a repetition in which every note was twice as long. Two questions were asked: Which dance tune was for the little girl and which was for the grandfather? Was the grandfather's tune the same as the little girl's, or was it different? Four out of the eight of the five-year-old children were able to conserve the melody, while four said the tune had changed. Seven out of the eight eight-year-old children succeeded in conserving the melody, two saying that the first part was 'twice as fast as the ending'. This kind of answer showed that they were trying to solve the task operationally by comparing tempi. The five-year-old group did not exhibit this kind of controlled judgment.

Piaget found that children under six or seven do not grasp relationships between parts of complex figures. They perceive a global picture of the total shape of the parts but fail to note the manner in which the parts are fitted together. To test the child's ability to grasp the relationship between the tones of a three-note figure, a melody was transposed to different pitches. After listening to the pattern played twice, the investigator and the child sang the melody and used the hand to indicate pitch levels. The child then had to identify which 'make-believe' child played the pattern 'incorrectly' (i.e. at a different pitch). The responses of the five-year-old children represented an intuitive phase of conservation. They could perceive the direction and contour of the patterns, but failed to observe the relationships between the intervals. Even the eight-year-old children's perception centred on the directional contour of the melodic line. The percentage of their replies that were correct was only 75%, compared with 63% obtained by the five-year-olds.

In the next task (see example 1), the five-year-old children had difficulty in conserving the melody when the rhythmic pattern changed.

Their perception centred on the rhythm. At the eight-year-old level, three children still found difficulty. 80% of their responses were correct, compared with 68% at five. Finally, a melody was repeated four times with various accompaniments. In one version the melody was incorrect. Two of the five-year-old group were able to select the third performance as being the incorrect one at the first hearing. Only two realised that the melody remained the same when an elaborate accompaniment was played. Perceptual concentrations upon the general effect of the total setting did not allow for a consideration of the melody *per se*. No stages in growth of ability to conserve melody with different accompaniments were evident. The score at eight (56%) was little better than the 50% at five. This was the only task to show hardly any differentiation between the two groups. More of the children might have given correct answers if the wrongly placed melody had been played before the misleading elaborate harmonisation. But Pflederer's result accords the general view (see below) that appreciation of harmony develops relatively late.

Pflederer concludes from her research that the answers at the five-year-old level were indicative of pre-operational thought, while those of the eight-year-old group had reached the intermediate stage of conservation, with a few individuals showing the beginning of operational thought in some of the tasks. The stages of conservation seemed easier to identify in the area of rhythm than of pitch. Her results were based on a rather small number of subjects and of items in each task. But a major research project is now proceeding in which a revised form of the tasks is being given to four age groups: five, seven, nine and thirteen (Pflederer and Sechrest, 1967).

Some of the tasks used by Pflederer have been applied before without reference to Piaget's ideas. Discrimination between duple and triple metre is often required of music examination candidates, they being asked to beat time to music played by the examiner. Mainwaring's rhythm tests (1931) are largely concerned with this ability. Lundin (1949) included a test of melodic transposition in his battery intended for older subjects. It was a test which his music student subjects found easy, as evidenced by the low percentile rank of twenty-five being allotted for only four errors. But such tests were mainly intended for group application. Pflederer, moreover, is concerned not so much with the child's immediate aural perception as with trying to discover from his explanations how a child thinks about what he is hearing.

Pflederer mentions the possibility of longitudinal studies of individual children to determine more accurately stages of conservation. These would be valuable, because of the variations in ability among children of the same age.

A striking example of marked individual differences of ability among children of similar age is shown in the following table of results obtained by Mainwaring with his tests, described on page 279:

Subjects	Pitch %age correct Mean	Pitch %age correct Range	Rhythm %age correct Mean	Rhythm %age correct Range
52 children, mean				
age 10·6: 29 boys	73·0	43·5 to 99·2	42·4	27·7 to 72·2
23 girls	84·8	65·8 to 100·0	73·6	50·0 to 94·4
31 girls mean age 9·6	64·4	30·1 to 89·7	59·6	27·7 to 83·3
34 boys mean age 11·5	80·8	51·9 to 96·1	61·6	5·8 to 82·3

The girls in the first group had received ordinary school music training for three years, but the boys had had no special music training at all.

Mainwaring was surprised by the low level of ability he found. For example, only 24 out of 83 of the first two groups gave correct answers to the 16 pitch items where they were required to state only whether or not the second note was the same or different from the first. Twenty-two children judged notes a semitone apart in the normal vocal range to be the same in pitch. Mainwaring also studied 83 boys whose ages ranged from seven to fourteen. Both rhythm and pitch tests showed definite tendencies to rise after the age of 9, but after 12, performance on the rhythmic tests declined. Mainwaring thought that this deterioration might be connected with adolescent clumsiness, especially in matters involving motor-co-ordination. On his tests of auditory recall, a more consistent tendency to rise with age was apparent. Between 12 and 14, the average score rose steeply from 42·4 to 62·3%.

Since they were tested in groups, his subjects did not have the advantage of being able to clap or sing, as did Pflederer's. In his article, Mainwaring mentions an interesting example of a boy aged five to six with an IQ of 120, who could march in time, modify his pace to changes in the music, and showed pleasure in rhythmic tunes. But even after he had learned to beat time to a tune in duple time played with an exaggerated accent on the first beat of the bar, he would cease to beat through a sustained note, vaguely increase pace during a series of half-beats, and find it impossible to regain the beat once he had lost it.

According to Gilbert (1893), a colleague of Seashore in his early days at Yale, a child improves in pitch discrimination twice as fast from six to nine as he does in the years from nine to nineteen. This conclusion was based on testing ten children at each age from six to nineteen with a pitch pipe adjustable to fine differences. Each child was given ten trials. On average the children of six could discriminate sounds three-eighths of a tone apart, at seven sounds less than one-third of a tone could be discriminated, at eight, one-quarter tones, and at nine, approximately one-eighth tones. The ten-year-olds had less fine discrimination than the nine-year-olds. After ten, discrimination improved, till at fourteen, sounds that were just over one-tenth of a tone apart could be distinguished. All but three of the children could, on the average, discriminate a semitone. The decline in discrimination at ten may well have been due to the smallness of Gilbert's sample, since it was not confirmed in Arnold Bentley's results with much larger numbers of children.

Arnold Bentley's tests (1966) arose from his interest in the musical development of younger children. He found that pitch discrimination improved between seven and fourteen by about 30% for 26, 12 and 6 cps differences, but by only 10% at the 3 cps difference level. His general conclusion was that the majority of children including the seven-year-olds, can discriminate a pitch difference of a quarter tone (12 cps at 440 cps) correctly, and that about half of ten- and eleven-year-olds and the majority of the twelve-year-olds and older can judge one-eighth tone accurately. Pratt (1928) found that adult observers could discriminate differences of one-eighth tone about half the time. Differences of about a quarter tone could be discriminated all the time.

With his memory tests, Bentley found that at all ages between eight and thirteen, there were some children who could score full marks and others who scored only one mark out of ten, or none at all. The steepest increase in mean scores was between the ages of eight and nine, where tonal memory scores improved by 15% and rhythmic memory by 16%. After nine tonal memory increased fairly steadily by an average of 6% a year till fourteen, while rhythmic memory increased rather less steadily, but by an average of 5·5%. On his own memory test, Wing

(1941*a*) noted that between eight and eleven memory for the immediate recall of tunes a few notes long develops.

The yearly increase on Bentley's chord analysis test was very small before 10. It was not till the age of 11 was reached that the mean score clearly exceeded the theoretical guessing score for the test. Between 11 and 14 the yearly improvement of the scores with age ranged from 5% to 8·5%.

Improvement with age was also evident in the results of an extensive longitudinal study carried out by Robert Petzold (1966). He studied the auditory perception of musical sounds on a total of over 500 children in the first six grades of Wisconsin Schools. He constructed a 45-item test by analysing a large number of songs used in the schools in order to find common tonal patterns. The patterns were chosen to represent:

1. Ascending Scales or Chords.
2. Descending Scales or Chords.
3. Ascending and Descending Scales and Chords.
4. Disjunct patterns.

These were tape-recorded with an interval of silence after each, during which the child tried to sing back the tune. The child's responses were recorded and later analysed. Petzold also used a phrase test in which the child listened to a four-bar phrase played twice and then tried to sing it back. The process of listening and responding was repeated till either the child was able to sing the phrase correctly two trials in succession, or till after ten presentations of the phrase, even though the perfect trial had not been achieved. Petzold found a steady improvement with age which tended to level off among the older children. The most significant changes occurred between the ages of six and seven. The phrase test was found to be very difficult; and even at the sixth grade (age 12) only one-third of the children managed to learn the phrase in 10 trials. Only 8 out of 90 children were capable of learning the phrase by grade four and retaining the scale for subsequent years. Furthermore, a second phrase given to these children in the sixth grade showed that they performed at the third grade level of competence, indicating that the learning process had not changed significantly during four years despite experience of a task of this kind. Children with low or high scores in the initial year continued to earn high or low scores during subsequent years.

Petzold also carried out a one-year pilot study of rhythm with 360 children. The children were asked to tap and to sing back common rhythmic patterns. They were also required to tap at a speed established by a metronome, both while the metronome continued to beat, and also over periods during which it was silent. The ability to respond accurately to rhythmic patterns and to maintain a steady beat did not change substantially after the age of nine. Maintaining a steady beat at tempo of 92 and 60 beats per minute was more difficult than at a faster tempo.

Stamback (1960) found a similar development between six and nine in children's ability to tap the rhythmic patterns of her test and to understand the symbols of test 3. Six-year-old children seemed to be able to deal with a limited succession of taps, but could not integrate several subgroups into a whole. After the age of eight hardly any child taps irregularly. Children of around nine or ten could successfully reproduce all but the most difficult items in her tapping test. She suspected that an important change took place around the age of six and that five-year-olds would all fail. Williams (Williams *et al.*, 1933) asked children aged from three to five to tap in time with the clicks on a Seashore rhythm machine. In the simplest form of the test, regular clicks at intervals of ·5 of a second, approximately 75% of the three-year-olds failed, only 25% of five-year-olds failed, at six there were only 4% of failures.

DEVELOPMENT OF HARMONIC SKILLS

It is generally held that most young children have no great appreciation of harmony, finding 'every harmonic accompaniment equally good, whether consonant or dissonant'. Valentine (1962) concluded from experiments on some 200 elementary school children, aged 6 to 14, that

> '1. No appreciable preference for concords before discords is discernible before the (average) age of 9, but at this age a marked advance takes place. . . .
>
> 2. It is not till we reach the age of 11 that we find the discords show a negative score, i.e. are on the average more displeasing than pleasing to the children.' [The negative score for discords at eleven was due to the more musical pupils, as measured by Stumpf's tests (see p. 25).]
>
> '3. At 12 or 13 we suddenly find changes which result in an order of preference for the various intervals, which is very similar to that given by adults.'

However, Valentine found that girls at a preparatory school were often three years in advance of the elementary school children in their judgments of intervals.

Rupp (1915) played a melody in E major to five children, three times, with the bass accompaniment in the keys of E, D, and F. The children accepted without hesitation the basses in the wrong key, except in two places where very strong discords were produced. Brehmer (1925) investigated the reactions of children aged 6 to 13 to various false playings of the scale. Only 20% of the 6-year-olds, but all the 12- and 13-year-old children, noticed a flattened seventh. Franklin (1956) concluded from such researches that the establishment of an ear for tonality is preceded by a stage where one can speak of an ear for pitch and an ear for consonance and dissonance, but not for simultaneous horizontal and vertical listening. An appreciation of harmony does not develop till the child is capable of attending to a bass as well as to a tune. Raven reports a rather similar finding on the development of ability to perform his Matrices test, which requires the subject to choose the one piece out of six that will complete a visual pattern. The child of five, Raven states, is sometimes satisfied if the piece he chooses to insert on the form board completes the pattern correctly in one direction only. Later, he begins to choose a piece that completes the pattern in two directions simultaneously.

Wing found that only by the age of 11 could children of average musical ability give answers to his harmony test that were much above the level of chance. Between 8 and 11, however, they could detect whether a chord had two or three notes, and whether the treble moves up or down. They seemed to develop some sensitivity to the effects of harmony before eleven, as shown by the results of Wing's cadence and discord tests. (These were two of the tests that were included in his early experiments.) Valentine (1963) thought that discords might have a certain fascination for children, because they appear to give a greater body of sound than do concords.

DEVELOPMENT OF APPRECIATION

Average scores on Wing's test of the appreciation of rhythm do not exceed the chance level till 10. This level is not reached on the other three appreciation tests till the age of 11. This finding might be compared with M. D. Vernon's (1960) conclusion that the normal child does not fully understand pictures and interpret them as a whole till about that age. This does not, of course, mean that younger children within their limited cognitive powers do not enjoy the kind of aesthetic experience to be discussed in Chapter XXII. Burt (Burt *et al.*, 1945), indeed, found that some of the youngest children (under eight) came very near to the judgment of experts on a picture ranking test and concluded that there might be some truth in the idea that as people grow up their artistic vision declines, i.e. they see what they know to be there, not what is there to be seen.

VIII

Music in Adolescence

As noted above, the average child's ability to appreciate the finer points of music is likely to be reaching a measurable level by 11 or 12.

Scores on musical ability tests go on increasing with age up to about 17 in the case of Wing's tests, but up to 20–1 years in the case of the Drake Memory test, with 'non-music' students. With music students, the improvement with age on Drake's test does not end till the age of 23. The difference is quite substantial; for the student who made twenty errors at 19 years would attain on PR 50, while the same score at 23 would be worth only PR 23. The norms for the music groups were based on altogether some 350 college students, not so small a sample as to make its representativeness suspect. An alternative explanation might be that the memory test was susceptible to the effects of the intensive training which presumably the older music students have received.

Meissner (see Schoen, 1940), who studied the mistakes which children between the years 8 and 14 made when they were asked to sing back tunes, found a marked improvement in the melodic memory at about the age of 13 or 14. Children of this age try to sing with expression and to produce notes of good tone quality. Meissner believed puberty to be a period when an interest in singing as an expression of mood and feeling develops. This deepening of the emotional appeal of music is perhaps the most noteworthy factor in music at adolescence. In many children it may take the form of an increased interest in popular music.

INCREASED LIKING FOR POPULAR MUSIC

Rogers (1956) played 57 pairs of recordings to 635 children aged 10, 13, 15 and 18. The two excerpts, lasting 45 seconds each, were taken from serious classical, popular classical, 'dinner' and popular music. The subject was simply asked which of the two he preferred. The results showed an overwhelming preference for popular music at all grade levels. Moreover, popular music was chosen to an even greater extent with increasing age. Upper-class children tended to like classical music more than did children from homes of lower socio-economic status. However, with age all the children seemed to conform more and more to a single pattern of musical preferences.

Wing (1941*a*), too, notes a liking for jazz and dance music among adolescents. In his thesis (1941*a*), he listed a number of reasons:

1. For the young child the appeal of music is derived largely from pleasant aural sensations and in this respect the classical songs learnt at school are superior to more popular forms of music. However, in adolescence, a connection between music and emotion begins to develop. It is the errand boy, not the young child, who sings or whistles to express his feelings. The words of popular songs usually centre on some primitive (often sexual) desire or express boisterous merriment, i.e. the sort of feelings that the adolescent can understand and share. The classical composer, on the other hand, expresses the emotional and intellectual outlook of the adult.

2. The adolescent is able to perceive and enjoy devices like syncopation, but is not so used to them as to be critical when they are repeated too often.

3. He is going through a stage of tearing himself away from dependence on the older generation. That popular songs are different from those he has had to learn at school and are to some extent frowned on by his elders, make them all the more attractive to him.

4. Popular dance tunes are easier to play when their simple tricks of style have been mastered. Each new song is merely a different melody with a similar harmonic background to the previous ones. The melody is usually uncomplicated, with one or two short and slightly varied phrases used in its construction. The boy with ability but little musical training can pick them out by ear without much trouble. A classic, on the other hand, is individual; each new piece has to be studied afresh. In addition, the technical difficulties in playing are often considerable.

Rogers found a sex difference that seemed to confirm the part sex plays in the increased interest of adolescents in popular music. The girls of 13 and 15 preferred popular music to a much greater extent than did boys of similar age. In his opinion, this was due to the earlier sexual maturation of the girls, for whom popular music had taken on a new social meaning. At the age of 18 there was no difference between the sexes in their preferences, nor was there a difference at ten.

We may question, however, whether asking children to choose between a very brief excerpt from a classical and from a more popular piece, is really a valid way of estimating their musical preferences. For one thing, their replies may merely conform to what they believe are the tastes of their friends of their own age. In any case, an expressed preference for popular music is by no means incompatible with an enjoyment of Beethoven or Handel.

SEX DIFFERENCES

Looking now in some detail at the similarities and differences between the sexes reported in various studies, we find that, on the whole, male subjects tend to make about the same scores as females on musical ability tests. Thus, in the Seashore test manual, it is reported that the 'sex differences were found to be very small and inconsistent from one level to another. Combined sex norms were, therefore, formulated.' No significant differences between the sexes were found on the Drake rhythm test, while on the Musical Memory test, there was a slight superiority for girls – amounting to about $1\frac{1}{2}$ points for either form of the test alone. The difference was too small to show in the norms. Wing obtained approximately the same average marks for both sexes at each age (but see below), so that adjustments for sex were not required. Such differences as Gordon (1965) found with his tests were too small to be of practical educational significance.

The table on p. 89 summarises the differences on the K–D tests, as described by Kwalwasser. These differences are quite small. Larger differences were found by Gilbert (1942), who tested 1,000 college students with the K–D tests: the mean for 500 women = 208 and for 500 men = 202·5 (out of a maximum score of 275). However, the female group had received about three times as much instrumental training as the male group. He attributed the superior scores of the women to the greater amount of their training and not to innate superiority because:

(*a*) when only the untrained members of the two sexes were considered, the difference disappeared; and

(*b*) when the subtests most susceptible to training were taken out, the differences were greatly reduced (and in the case of the untrained subjects, actually reversed).

Though Wing found that no adjustment of sex differences was required to the norms for his battery, he reported that after the age of 14, girls seemed rather better than boys at the appreciation tests, though the two were still equal in performing the ear acuity tests. The difference amounted to about 4 out of 80 marks. Discussing in his thesis possible reasons for the difference, Wing rejected both errors of sampling and differences of training, since the children came from similar districts and homes. Moreover, the same phenomenon was noticed in a co-educational school, where the two sexes had the same teaching in mixed

TABLE VIII.I *Sex differences on performance of the K–D tests*

Subjects	Males superior	Females superior	$\bar{x}$ Male score	$\bar{x}$ Female score
4,247 American children grades 4 to 12	Time, intensity* and quality discrimination	Rest of K–D test	179·70	182·05
3,588 boys, 3,833 girls grades 4 to 12			179·05	181·25
500 University students	Quality, intensity and time and pitch discrimination	Rhythm, melodic taste, tonal memory, tonal movement, pitch and rhythm imagery	197·31	202·54
700 Junior high school children			178·30	179·40
6,000 European children	By 20 points in Hungary	By 14 points in both England and Italy	191·06	191·49

* According to Kwalwasser's text but not according to his figures.

classes, and where they were of similar IQ and home environment. Furthermore, a difference of training or of interest could be expected to have at least as great an effect on the first three tests. Two possibilities which, Wing suggested, might repay further research were the change in the boy's attitude to music which might accompany the physiological changes in the larynx when the boy's voice breaks, and the supposition that girls have a greater preponderance of introverts after adolescence and 'that introverts make better listeners to music which requires appreciation'. Another possible explanation might be in the earlier puberty of girls, i.e. girls are merely more advanced in their appreciation of music than are boys.

Very briefly, Wing's argument on the connection between introversion and women's interest in listening runs as follows: introverts tend to find their chief delight in music as listeners (or as composers if they have creative talent). Women are, on the whole, more introverted than men (Wing quotes Jung's statement to this effect). Therefore, women can be expected to make better listeners than men. As Wing points out, it is obvious from a visit to any concert that women do form the greater percentage of concert audiences. This cannot be ascribed only to the fact that they have more free time, for it is found that the percentage of males in an amateur orchestra is far higher than at a concert. Wing attributes the larger percentage of amateur male performers to a personality, rather than a musical, difference, since men are more extroverted than women and musical extroverts tend to gain more enjoyment from performance than from listening. Wing, like Jung, recognises, of course, that most people are of mixed personality and that with musicians 'there will be no sharp demarcations'.

Some evidence that seemed to support Wing's view emerged from Shuter's (1964) factorial study of 100 male and 100 female Training College students. The aim was to compare a group of men with a group of women who were of similar age, general ability and musical level. (Intellectually the women may have been slightly superior, since more of the better qualified men go to University.) The scores of the subjects included ranged from 78 to 110. (Maximum score is 136. The mean for persons of 17 or above is 76.) This higher-than-average level of ability was chosen, as being the one at which the Wing tests are most reliable and valid. The mean of the men's scores was 91·05 (SD 8·6), that of the women's, 91·25 (SD 8·57). (Before selection, the average for the women was possibly rather higher than for the men. At least, Shuter had the impression that she had to work through a somewhat larger number of the men's results to find 100 above average scores.)

Subtest results showed that the men did the pitch test rather better than the women with a similar average score for the whole test, while the women found the last three appreciation tests easier than did the men.

With both the men and the women, a broad factor of general musical ability was found, but with the following interesting difference: appreciation of changes of intensity appeared to be lacking in the main factor for the men (see Appendix II). Appreciation of phrasing and of rhythm seemed to have a much more prominent place in their musical ability, while appreciation of the appropriateness of intensity changes appeared to be much stronger among the women. Appreciation of rhythm would seem to play a negligible part in the women's musical ability, as evidenced by the zero loading of this test in the general factor. It is, perhaps, not too far-fetched to consider this difference as connected with the difference which Wing postulated. The extrovert is likely to be interested in, and particularly sensitive to the rhythmic movement of music. If boys tend to be more extroverted in their approach to music, rhythm might be more likely to appeal to them. If girls are more introverted, more interested in music as a means of expressing feeling, it might well follow that appreciation of changes in intensity would be a more important component of their musical ability. It is, after all, by changes in intensity, indicated by 'expression marks', that emotion is most obviously made overt in music. Also, Wing found from the introspections of many of his subjects that it was his intensity test that they were most inclined to solve by 'intuition' rather than by analytical judgment. It might be interesting to investigate whether persons high on the extroversion side of an objective personality scale would be more successful with the Rhythm test and those who score high as introverts on the intensity test, irrespective of the sex of the subject.

The greater appreciation of phrasing found among the male group agrees well with the view that boys are usually more interested than girls in how musical compositions are constructed, since it is through phrasing that the construction of a movement is made clear. As noted below, the achievements of men in creative activities remain much greater than those of women.

There may, then, be some qualitative differences underlying the differences in scores between the sexes which are quantitatively quite small. As Allport (1937) remarked on judgments of personality, 'All in all, there are plenty of reasons to account for woman's superiority . . . The wonder is that their superiority over men is not more marked than it is.'

Valentine (1962) reported that the introspective remarks of 146 subjects appeared to show that there were more men than women who were deeply sensitive to the impressions of musical intervals. Moreover, the proportion of judgments of high aesthetic judgment value made by 52 men was higher than among 84 women. This might be partly due to Valentine rating 'subjective' judgments low. (He rated perception of the 'character' of intervals higher than musical 'association' of the interval played with a following one imagined by the listener.) Possibly, too, women might find the task of judging single intervals lacking in emotional appeal or musical interest. This (as Farnsworth, 1931, implied) might also explain why sex differences are sometimes found in the performance of tests like the Seashore. Such tests lack 'musical' appeal, but boys might find them of scientific interest. As far as interest in music is concerned, a much wider difference (5 to 8 points compared with 2 or 3) separates the sexes on Gaston's (1958) interest inventory than is apparent in the norms for the tonal items of his test.

At the musical ability level of the ordinary schoolchild, girls appear to do better than boys both in musical knowledge and in class singing. Reporting the results of testing over 4,000 children with the Kwalwasser–Ruch Musical Achievement Test, Kwalwasser stated in his book that the fourth-grade (age approximately 10) boys were only 5 points behind the girls, but by the eighth grade their mean score was lower by nearly 30 points (equal to two years). Semeonoff (1940) found, however, boys achieved rather higher scores on his adaptation of the Kwalwasser Music Information and Appreciation Test than did girls. Boys were significantly better at selecting words which gave the best description of a composition; girls at preferring the acknowledged masterpiece rather than two pieces of less musical worth. In the validity experiments on the K–D tests referred to on p. 31, a much higher proportion of boys was found among children judged as poorest at singing by their teachers and a much lower proportion of boys were among children selected as the best singers in the class. Arnold Bentley (1963) found that the percentage of 'monotones' among boys was much higher, even at the age of seven and it failed to decrease with age as much as did that of girls (see also p. 136).

Petzold (1966) concluded that the differences between boys and girls appeared to be related to the nature of a musical task. With two out of three of his groups, the 45-Item Test showed boy/girl differences that were highly significant. However there were no significant differences between boys and girls for the harmony or rhythm studies or for the phrase test of his longitudinal study. He attributed these differences particularly for the older children more to attitude and motivation than to basic differences of musical competence. The girls generally continued to improve their performance while the boys showed either only slight improvement or did worse in grade six than in grade five. Petzold believed that this was partly related to the attitude of the boys towards using the singing voice – they lacked both confidence and competence in being able to view singing as a natural musical response. We must remember too that these boys were still well below the age at which their voices would break.

At higher levels of achievement, far fewer women than men win distinction. A small number of Sergeant's ISM group were considered by him to show particular gifts because of their internationally accepted reputation. Members of this specially gifted group included Sir Adrian

Boult, Gerald Moore and Clifford Curzon. Only $12\frac{1}{2}\%$ of this highly talented group were women. For instrumentalists, this may be partly due to the prejudice against women players, which seems to be dying only gradually, judging by the small number of women among the permanent players in the major European orchestras.

But, if a woman of talent encounters prejudice when she seeks employment in an orchestra, there would seem to be no such barrier to prevent her composing. (She might, however, experience more difficulty than a man at getting her works played.) It seems unlikely that women students whose compositions showed real merit would be refused tuition and encouragement. Yet in the spring of 1965 a concert of works entirely by women composers was a novelty in London. Dame Ethel Smyth, Elizabeth Lutyens and Thea Musgrave are among the very few women to achieve eminence as composers. There may be some genuine sex difference here. Perhaps the highest level of musical genius occurs much more rarely among women, as appears to be the case in literature and painting also. However, there does not seem to be in the musical world even the equivalent of the numerous women novelists who are usually competent and sometimes highly distinguished.

IX

Precocious and Unusual Talents

One of the characteristics of musical ability is its tendency to emerge at a relatively early age, even among not specially talented children, not to mention the outstanding gifts displayed by many quite young performers.

From data collected by Haecker and Ziehen from 441 cases, Revesz (1953) concluded that nearly half of the children revealed musical aptitude between their second and sixth years. From parents' reports on the age at which various abilities were first noted in his sample of gifted children, Terman (1925) found that, except for general intelligence, musical ability was shown at the lowest age. The average for boys (n. = 91) was 4·6 years, for girls (n. = 108) 5 to 16 years. The reliability of parents' reports may be thought to vary considerably, but at least the public performances of musical prodigies provide reliable evidence of precocity.

We have already mentioned the early ages at which the professional musicians who filled in Sergeant's questionnaires commenced their musical training.

Among the virtuosi instrumentalists studied by Amram Scheinfeld (see p. 116) musical talent appeared at an average age of four and three-quarters and in the case of the Juilliard music students at five and a half. For example, Yehudi Menuhin was studying the violin seriously at three and a half. In the case of the opera singers formal training did not begin till the age of 15½ for women and 16½ for men, though they showed musical talent about seven years earlier. The average age of their professional debut of the virtuosi (not merely their first public appearance) was 13¼.

The history of music provides many examples of conspicuous musical talent being displayed by young children, not only as performers, but also as composers. In the course of her descriptions of the early lives of eleven eminent composers, Cox (1926) mentions many instances of their precocity.

For instance, at the age of 12 Beethoven was already able to read and play the most difficult and involved scores at first sight. Three sonatas for the piano were published before he was 13. His first composition is said to have been written at 10. A funeral cantata in memory of the deceased English Ambassador, this piece apparently aroused great astonishment at its originality. Handel was producing a church cantata every week when he was only 10. The famous story of his playing a clavichord in the attic at night is associated with the age of five or six. Haydn is said to have begun composing at five. A relative who was a schoolmaster and choir leader was so impressed by signs of the child's musical talent that he offered to adopt him and undertake his musical education. At the age of six Haydn could sing several masses in the Church choir and play a little on the piano and violin.

Mendelssohn's creative gifts developed rapidly and prolifically after the age of ten. In his eleventh and twelfth years he wrote some 50 or 60 songs and pieces for piano, violin and organ. Benjamin Britten is an example of a present-day composer who began to write music at five. By the age of 15, he had over 100 compositions to his credit.

Mozart was one of the greatest (if not the greatest) musical prodigies of all. His talent was discovered because of the keen interest he took at the age of three in his sister's music lessons. He would amuse himself for hours picking out thirds and showed a good memory for tunes that he had heard.

His father, himself a talented musician, soon started giving Wolfgang lessons and writing down the little pieces which his son composed. In 1762 he took the children on their first tour. By the age of seven Mozart was giving public recitals in London. At this time too his outstanding sense of absolute pitch was discovered, as well as his remarkable skill with the violin and the organ which he had never been taught. His repertoire included the naming of any note, improvising in any key and transposing to any key. His famous feat of memorising and writing down the Allegri *Miserere* after hearing it performed in the Sistine Chapel dates from the age of 14.

Pointing out that it is difficult to assess the amount of Mozart's early training and the effect of his musical environment, Richet (1900) reported the case of Pepito Areola whose talent appeared before he had had any training whatsoever. Pepito's father was not musical but his mother had played the piano at the age of five, while his maternal grandmother is reported to have been a good guitar player. When hardly two and a half years old, Pepito played tunes on the piano. Sometimes they were tunes his mother had played or sung, sometimes they were of his own invention. When investigated at the age of 3:7 by Richet he could play 20 pieces from memory, including the harmonies. He used clever fingering to make up for the smallness of his hands. His improvisations showed some feeling for form. He appears to have been rather temperamental and resented being corrected. He refused to play on any instrument except his mother's piano. When Pepito was 6:2, Stumpf found that he could sing any note he was asked for and name any note played on the piano or even on an unfamiliar instrument.

The most detailed investigation of a musical prodigy was made by Revesz (1925) who was able to observe Erwin Nyiregyhazy from his sixth to his twelfth year. Erwin's father and paternal grandfather were singers in the chorus of the Royal Opera in Budapest. His mother also possessed considerable musical talent. His younger brother, was 'remarkable for a strong feeling of rhythm and a very good musical memory'. According to his father, Erwin tried to imitate singing before he was one year old. In his second year of life he would reproduce correctly melodies sung to him. At the beginning of his fourth year he began to play on the piano everything he heard. He also improvised. In fact, at 3:6 he had already composed little melodies. From his fifth year he received some piano lessons, but regular tuition in music began only when he entered the Academy of Music at six. When tested by Revesz, his musical ear at seven was already developed to an extraordinary degree. He had absolute pitch, being able both to name the notes played to him and to sing any note on request. He could analyse complicated chords with greater accuracy than David Popper, the cellist tested by Stumpf. His immediate musical memory was nearly as good as that of an adult pianist whose musical memory was known to be very good and his power of retention for a half an hour or a day was much better. He could memorise melodies harmonised in a simple manner with the greatest ease and had no difficulty in retaining a great number of operatic airs in his memory, but could usually reproduce without mistake only the melodies without the harmonies. He reproduced faultlessly at the third attempt a 13-note tune played to him by Revesz. Two years later he was able to reproduce without mistake a five-bar theme at the second attempt, the time taken for learning being 22 seconds.

Although both Nyiregyhazy and Areola became professional musicians, neither seems to have achieved the outstanding renown that might have been predicted. According to Nicolas Slonimsky (1948), himself a former child musician, only about 10% of child prodigies become adult virtuosi.

Speaking from the experience of having taught over 3,000 pupils, Cortot (1935) stated that the proficiency which some children display is no more than the manifestation of dexterity and an extraordinary natural imitative faculty. However, many infant prodigies, like Yehudi Menuhin and Lorin Maazel, do become highly esteemed adult musicians. 70% of the great violinists listed in Leahy's *Famous Violinists* were prodigies (Drake, 1957, p. 13).

Certain individuals have adopted careers outside music and later earned renown as composers. Borodin was a professor of chemistry. Moussorgsky and Cui had military careers, Rimsky-Korsakov was a naval officer for nine years. All, however, had shown aptitude for music early in life. Borodin was able at eight to play the piano by ear. He was taught to play the flute and the piano. Moussorgsky's mother gave him his first piano lessons. At seven he could play small pieces by Liszt and before nine a Field concerto. Cui showed a precocious talent for music and was taught the piano at an early age (*Grove's Dictionary*, 1954). Soon after beginning the piano at six, Rimsky-Korsakov could name any note played to him (Teplov, 1966, p. 168).

This does not mean that training for music must begin at such early ages. For example, Malcolm Tillis had sufficient talent to become a viola player with the Hallé Orchestra. His family was not particularly musical. At 11 he was taken to see *Carmen* and fell under the spell of music. After that he took every opportunity of listening to music on the radio, but did not begin to learn an instrument till he was 15 (Tillis, 1960).

Leonard Bernstein had no opportunity to learn a musical instrument till his family acquired a decrepit piano when he was 10. Even when his exceptional talent became apparent, his father tried to discourage him from taking up music as a career. Henry Cowell, the American composer, had had no musical training at all before he was 14. When he was 15 and living in the direst poverty, he scraped together enough money to buy an old piano and taught himself to play it. By the time he was 17 friends recognising his talent subscribed to a fund to enable him to receive a proper musical training (Burks, Jensen and Terman, 1930). Whether the talents of these musicians might have developed even more if they had had opportunities earlier in life is something that we cannot determine. Cowell himself thought than an orthodox musical training would have hindered rather than fostered his creative ability. The lack of a musical instrument of any kind made him turn to an inner world of music of his own creation in which he sought for effects not heard in traditional types of musical compositions (see Chapter XXII).

THE IDIOT SAVANT

The so-called 'idiot savant' is an individual of very low intelligence who shows some special aptitude, such as musical ability.

The typical musical accomplishments of idiot savants show some similarities to those of infant prodigies. They can often play a tune by ear after hearing it only once or twice. Some learn to read music. But their lack of general ability and of emotional stability prevents them from developing their talent normally.

A well known case was that of Blind Tom (Drake, 1940), who became a vaudeville artist. With no more intelligence than a child of six, he was evidently able to memorise a piece from one hearing, and to play two tunes and sing a third at the same time. Afterwards he would join the audience in applauding himself. His manager may, of course, have encouraged such signs of 'idiocy' for publicity reasons.

Tredgold (1922) mentioned a woman at the Salpêtrière Institution. Though an imbecile, a cripple and blind from birth, she could sing any selection of tunes which she had heard. Her fellow inmates came to her to have their mistakes in singing corrected.

Rife and Snyder (1931) described a blind imbecile girl who could play a new and difficult piece on the piano after hearing it only once. A musician visiting the Vineland Institution where she lived asked her to play an unpublished composition of his. She was able to do this perfectly after hearing it only twice.

Though according to Tredgold, the special talent shown by idiot savants has rarely been marked in their ancestors, Rife and Snyder decided to study as many of the relatives as possible of such cases as they could locate. By addressing an enquiry to 55 American institutions for the feeble minded they succeeded in finding 33 idiot savants, of whom eight showed a special talent for music.

They studied personally a case earlier reported by Minogue (1923). *XY* had developed normally till three years old and as soon as he could talk had learned little songs. After contracting spinal meningitis he was left mentally impaired. At the age of 14 when he entered Letchworth Village his IQ was 62, at 23 (when Minogue described his case) it had fallen to 46. He had remarkable pitch discrimination and an unusually good tonal memory, being able to play jazz or classical music by sight or by ear. He sight-read the *Marche Grotesque* of Sinding and an accompaniment for a singer. Though emotionally unstable, he played well when willing to attend, but produced no original compositions and was unable to learn to dance. When a child he had received two years piano tuition but had been abusive to his teacher. His memory for time, for places, and for events, as well as for any composition he had ever learnt, was described as 'almost phenomenal'. His parental grandmother and a cousin were said to be pianists of exceptional ability.

Rife and Snyder also mentioned an idiot, aged 35, who could play on the piano any tune sung to him. The chords he used were harmonically correct, although he had never received any training in music. Some of his normal brothers played or sang. Another idiot, aged 19, could play by ear anything he heard. He had a feeble-minded brother with no musical ability and a normal, though blind, sister who played the piano and composed.

Owens and Grimm (1941) reported the case of a woman of very low intelligence whose adult mental age was only two years nine months. She played the piano in the ward of the Minnesota Institution which she had entered at 14. She played by ear popular music heard on the radio, also hymns, but seemed to need an auditory stimulus she could copy. Thus she had little ability to play tunes named, but played Brahms' Lullaby in the key in which it was hummed to her. Her home had acquired a piano when she was five or six but she had never received any tuition. Two of her four sisters played the piano by ear. Her speech was very limited and indistinct.

A more detailed study of an idiot savant with musical talent was reported by Scheerer, Rothmann and Goldstein (1945). This boy was classified as an idiot savant by the New York Neurological Institute, though a psychiatric consultant believed he might be schizophrenic.

When tested by Scheerer at 11 and at 15 *L*'s IQ was around 50. He first showed signs of remarkable interest and ability in music, rhythm and counting in his third year. He could recognise a melody if only part of it was played. By his sixth year he knew the melodies and names of many compositions and had absolute pitch at least for the notes of the piano. As he grew older he would sit for hours at the piano playing monotonous sequences. In spite of his gift for music, he was unable to develop it actively and creatively since his learning ability was limited to an unreflective manipulation of the keyboard. He reluctantly learnt to read music, was unable to practise, and could not repeat musical sequences of his own on request. For several years he had an almost obsessional interest in an aria from Verdi's *Otello* which he never seemed to get tired of hearing. Without knowing Italian he could sing the words phonetically. At 12 he was taken to a musician who played a piano piece which *L* did not know. When asked to repeat it the boy did so, according to the musician amazingly well – the melody was correct and the accompaniment adequate. Among his other skills he could spell correctly many words both forwards and backwards. His span for the immediate recall of numbers was seven forwards and six backwards; he could also add up correctly the total of 12 two-digit numbers just as quickly as one could call them out. However, he would make errors in multiplying or subtracting larger numbers although he had received instruction in arithmetic. When introduced to a new person, he would ask the date of his birthday and then in a fraction of a minute be able to tell exactly which day of the week the birthday would fall in five years time. On the other hand, *L* failed miserably on any activity requiring abstract thought. Scheerer and his colleagues believed that his talent for music could not function normally because of this impaired abstract capacity. He spent so much time on music and counting because such activities were his only means of self expression and of being able to come to terms with his surroundings. Both his paternal grandparents were musical; his father was very quick at manipulating numbers. His mother, who was a former schoolteacher and not particularly musical, devoted her life to bringing up her son.

Anne Anastasi and Raymond Levee (1959) described the case of a high grade adult mental defective with exceptional musical talent. Levee had been employed as his tutor for two years. Both *S*'s parents and his brother were college graduates. Though reported to be normal at birth, while in hospital he contracted epidemic encephalitis which led to permanent brain damage. Before he could speak, *S* was able to hum tunes. In fact he was eventually taught to speak by a speech therapist through the medium of lyrics. His musical education began at the age of seven, under eminent teachers and concert pianists. His musical ability was judged to be outstanding by his teachers and by musicians who had played with him. At one time he played the piano at rehearsals for a leading chamber music orchestra. He was reported to be an excellent sight reader who could also play by ear when the occasion arose. When listening to music he usually assumed a critical attitude towards the performer and would not listen to anything but classical music played by experts. He regarded his own playing as serious work, practising from six to nine hours a day. Apart from his music he was lethargic and had only weak and short-lived affections.

S was also gifted with an outstanding rote memory. For instance, after a single silent reading of a two and a half-page article, about George Washington, he reproduced it verbatim. However, he was unable to grasp the significance of Washington's self-sacrifice described in the article. His memory for past events was equally conspicuous. If asked about something that had occurred a few days previously, he seemed at a loss to remember the details. Yet after a month or two he could give a full and accurate report of the circumstances and time. He acquired a large store of information about classical composers – where they lived, when they were born and died, when each composition was first published and played in public, etc.

Tredgold (1922) considered that the ability of idiot savants is 'the result either of some primary anomaly or of some fortuitous circumstance of early life which has aroused the child's interest and thence led to the concentration of all his mental activities upon one subject . . . the talent . . . certainly owes much of its development to constant exercise'. His son R. F. Tredgold (Tredgold and Soddy, 1956, p. 448) thought that it now 'seemed more likely the condition is more of emotional origin, in which intellectual development has become . . . obsessionally canalised'. While this may be partly true, some initial endowment of aural capacities would also have been necessary before the concentration of interest could result in actual music making.

It is, of course, very difficult to test a person of very low intelligence. Rothstein (Scheerer *et al.*, 1945) tried to use the Seashore tests on two adult mental defectives. One, who had a mental age of nine, could play the piano and had 'a good sense of rhythm, excellent pitch discrimination and tonal memory'. The other, with a mental age of only four, could reproduce tunes if they were predominantly rhythmic. Both failed on the Seashore music tests; despite simplifications they could not follow the directions and kept repeating the same answer irrespective of whether it was correct or not.

The writer tried to apply the Bentley music tests for young children to a few patients who had shown some musical ability in a hospital for subnormals. At the second attempt a girl who played the zither achieved a score of 17 out of 20 on the pitch test. A second patient who was very fond of music made a score of 11. A middle-aged woman (IQ 51) who had at one time played the piano by ear for the hospital Physical Training class scored only 4. None of these patients was able to understand the requirements of the other Bentley tests.

Mongols are sometimes said to have a good sense of rhythm. To test this notion, Blacketer-Simmonds (1953) matched 42 mongols with a similar number of non-mongol defectives. The subjects had to listen to three repetitions of three rhythmic patterns played on drums and to try to reproduce them. 18 mongols and 14 of the non-mongols were judged to show 'a good sense of rhythm'. Cantor and Girardeau (1959) tried to obtain a more quantitative measure. They asked their subjects to distinguish between two rates of metronome beats, 120 per minute and 88 per minute. They compared 44 mongols in a Tennessee hospital with 24 much younger normal children. The chronological age of the mongols averaged 12:4 and their mental age 4:4. They were significantly inferior at judging the rate of the metronome test. Unfortunately, the mental age of the normal children was 5:6 and this may have made some difference to the results.

X

Conclusions and Educational Implications

In Chapters VI to VIII, the development of musical ability from infancy to maturity has been traced. At every stage marked individual differences of ability are found. If the environment is at all favourable, signs of any strong talent that may exist in the child are likely to appear early in life. The course of development of musical perception seems to progress from an indefinite perception of a diffuse auditory field to definite perception of parts, which are later integrated into well-perceived wholes. At the first stage of melodic perception, the child can sing only the general shape of the tune. Later, with the development of a sense of tonality, the intervals between the notes are reproduced more correctly.

Interest in the variations of timbre and discrimination of differences in the loudness of sounds are likely to appear relatively early. Some observers think that capacity to deal with the melodic aspects of music precedes ability to deal with the rhythmic side; others held the reverse view. It is generally agreed that perception of harmony comes later. Appreciation of the finer points of music, as tested by the Wing appreciation tests, seems to develop later still.

There is some evidence that an interest in the sound of the individual note is an important factor in the development of absolute pitch. If the child begins to learn an instrument at the age when he has this interest, he may develop absolute pitch. If he is older when he begins to learn music, he may by then be more interested in other aspects of music and pay little attention to the sound of each note.

We would agree with Emett Wilson when he says that since music is an art dependent on feeling,

> as soon as we lose feeling, music is dead. Especially at the beginning the feeling must be one of fun, of pleasant exploration. As soon as the idea of *ought* or *must* comes in, the interest in playing and the development of musical taste is inhibited. Compulsory drill is dangerous unless and until the child – and, for that matter, the adult – feels the need of acquiring technique.

> Fortunately, encouraging the child's natural development is just the help the parent is best fitted to provide. . . . To be noticed, musical tones must bring some pleasant experience or they will not be noticed. The imitation of a cow mooing, the cry of *peek-a-boo*, the comfortable rhyming and the satisfying rhythms of *Mother Goose* – all develop a feeling for the fundamental elements of music.
>
> Singing to the child or with the child sharpens his consciousness of melodic units – melodic vocabulary. Since songs have rhythm, the rhythm units are also developed.

Even before the infant is ready to sing with someone else or with a record, parents can do much to aid his musical growth if they will imitate the child and sing back to him his own sounds. As the child grows he will add new sounds to his repertoire. The parents should progress *with not ahead of* the child (Brody 1953).

Shinichi Suzuki (Eastman School of Music, 1966) believes that the earlier the child can learn to play and listen to music, the better. Since 1954 more than 1,500 very young children trained by his method, have performed on the violin at annual festivals in the Tokyo Sports Palace. Their lessons begin when they are as young as two and a half or three years old, with the mother or father attending each lesson and actually learning with the child, taking notes, and learning the tuning and the correct posture. The parents are expected to help the child in his daily practice, encouraging him, but not forcing his progress. All learning is by memory; no printed music is used during the first two or more years, until his technique is well-established. All pupils using the Suzuki method follow a programme of carefully selected music.

Such a method of rote learning has been used for centuries among the gypsies of Hungary and Rumania. However no one has taught a large group of children in this manner before Suzuki. An investigation was begun in 1966 at the Eastman School of Music, to find out how far it would be possible to adapt such a method in America. In 1964 at the San Fernando Valley State College, California, Tibor Zelig (1967) started an experimental programme of violin instruction for very young children. At the first meeting the children were asked to name or sing the songs they knew. All seemed to know 'Mary had a Little Lamb', so they were shown how to play it by manipulating their fingers on the strings, and how to draw the bow across the strings. At first the playing positions adopted by some of the children were quite crude, but after six months they began to look like well-trained violinists. At the beginning, short individual lessons were given, later a mixture of individual and group playing was found to be effective. Some of the children learned solely by ear, while others found it simpler to learn the fingerings of a song from a numerical chart. Zelig concluded that the programme had proved worthwhile, even for children as young as three and a half.

Certainly for children who show exceptional talent, the sooner that instrumental lessons are begun the better, especially if there is any possibility of their becoming professional musicians.

For most children, however hearing and singing songs would seem to be the most natural approach to music before the age of five or six.

Martha Colby (1935) found that children between three and a half and four and a half all made some progress in learning to manipulate a tin fife. After a month all but two could produce most of the melodic patterns within a major third. But after that few of the children made substantial progress. She thought that a minimal amount of auditory-manipulative training might foster an interest in instrumental music and probably deepen its aesthetic value later on. However, she concluded that it would be better for children of this age to learn songs.

Brown (1936) investigated the optimum age for starting to learn an instrument such as the piano. He obtained a correlation of ·60 between age and rate of learning and concluded that the critical age for starting piano lessons was about seven. However, Mary Cochran (1930) claimed that children of five are ready to begin the piano, but only from teachers prepared to adopt a suitable approach for children so young. She stressed the importance of the kinaesthetic sense and evolved games that would make the child conscious of his movements. Gesell and Ilg (1946) found that the typical seven-year-olds among their sample of middle-class children will often express a strong desire to take piano lessons. They thought that the craving should be satisfied, so long as the teacher did not insist on practice between lessons. By eight, the child's initial interest may have died out, unless someone plays with him, or sits with him while he plays. He enjoys playing duets. It may be wise to interrupt lessons for a while till he is ready to return to them. If, at nine, his interest in taking lessons still persists, he can be expected to apply himself in earnest.

For amateur performance, it is not too late to take up an instrument in the teens or even later. In fact, William Cramer (1958) found that successful achievement in instrumental performance at the four to eight grade levels is significantly influenced by the motor development of the child and concluded maturational conditions for the beginning of purposeful study are optimum at grade seven (thirteen years old).

His conclusions were based on the results achieved on the Watkins Objective Measurement of Musical Performance Scale by 64 students, aged 10–14, after a year of class instruction on an orchestral instrument of their own choosing.

MUSIC IN SCHOOL

Pflederer concluded from her investigation that the overt interaction of the child with the musical problem seemed to be of primary importance in developing conservation. The child also needs models to imitate. The musical experiences provided by the curriculum should stimulate the maximum amount of growth at each stage. Musical experiences designed to clarify incipient concepts must precede attempts to intellectualise them. To sharpen the child's ability to discriminate between patterns and to follow thematic development of the patterns, experience with a large repertoire of these patterns in many and varied guises is needed.

Her recommendations are largely in agreement with those of Wing and Bentley. In his Ph.D. thesis, Wing outlined a school music course based on his own experiments. Arnold Bentley (1966) devotes the last chapter of his book to the practical application of the findings of his investigation.

For children between five and seven, Wing advocated a wide variety of songs, including those with considerable change of key. He questioned the value of placing too much emphasis on such activities as clapping to music, playing in percussion bands, and 'Music and Movement'. If children are to march in time, he suggests that they should sing at the same time. Bentley too believes that pitch is the most important factor, and that one good way for children to become acquainted with music is to listen to it, and wherever possible, to sing what they hear. However, rote-singing may soon become not sufficiently challenging – and, therefore, possibly boring, especially to the more musical children. He concluded from his research that the majority of children are ready for work requiring analytical judgments by the age of eight. Children of that age should be able to take in their stride the singing of diatonic and chromatic semitones; any difficulty that is encountered seems to come from naming the sounds, rather than from inability to discriminate between them. Being taught to name the sounds in their melodies helps them to come to grips with the music. In Bentley's view, by the age of 12, the optimum period for learning to name sounds may have passed. Neglect of this activity may, he believes, be responsible for lack of success in much primary school music – resulting in loss of interest among the older pupils.

Wing points out the importance of using staff to write down music, as well as to read it. It is, of course, all the more interesting if the children can write down their own compositions. Doig held classes on Saturday mornings in Cleveland, USA, where children were given the opportunity of setting words to music and of, for example, composing little valses. Some of their tunes were quite charming. The older children tended to produce rather more stereotyped compositions. It is perhaps a reflection on the difference in the place given to music as opposed to the vernacular in school work that this musical equivalent of routine oral class composition merited a series of three journal articles. Pflederer also recommends that the child should be encouraged to use his own original rhythm patterns in varied tonal contexts and his own tonal patterns in varied metre and rhythmic contexts, as an aid to the clarification of tonal and rhythmic relationships.

Both Wing and Bentley agree that the average child is hardly ready for two-part work before 11. Wing thought that a beginning might be made with the singing of rounds before then. This could later be followed by cadences and some discussion of the effects of discords. Bentley suggests the use of such instruments as dulcimers and chimebars as much easier than part-singing, as an introduction to polyphonic work. These help children to apprehend and analyse concurrent sounds, even though playing such instruments is usually an exercise in manipulation, rather than in tonal thinking.

In the years before 14, Wing thought it wise not to emphasise the emotional appeal of music. Attention should be directed rather to striking and unusual effects, and to the fine shapes of tunes. Children between 11 and 13 should find descriptive music easy to appreciate. Wing suggested that a rather different approach to music was needed when teaching boys than when teaching girls. Girls would gain a considerable amount of musical benefit and pleasure from advanced eurhythmics. Boys are likely to have a greater interest in the analytical and constructive aspect of music – how certain effects are produced, etc. Some discussion of chord construction, turning three or four notes into a melody, even simple harmonisations could be included in their curriculum. The making of instruments could lead to studying the historical development of the instruments of the orchestra and of the symphony, and to learning to identify instruments by ear. Playing the percussion parts with a recording of a symphony will give training both in score reading and listening to good music. With boys, the construction or analysis of melodies provided a simple approach to phrasing and to form. With girls, phrasing might be approached through eurhythmics.

The musical powers of children of over 14 are not far below those of adults. They can be expected to do a considerable amount of work as a leisure occupation, e.g. in preparing for concerts. Both as listeners and as performers they should be encouraged to form their own opinions, on the value of works and their interpretation. Analysis of sonatas and even acoustics (from the point of view of the playing and construction of instruments) are likely to interest boys. With girls, the structure of music might be better approached from the point of view of how it may contribute to the building up to emotional climaxes. The history of music is better studied through the history of composers, since girls are more interested in people than in the technical aspects of music.

An example of what can be achieved in music with adolescents that has received national publicity, is the playing and composing of the pupils of Cirencester Grammar School under Peter Maxwell Davies. Apart from any artistic worth that the music so produced may have, it is valuable in stimulating the pupils' interest in mastering the technical requirements of writing down and performing music.

PART III

Hereditary and Environmental Factors

XI

Methods of Genetic Study

GENERAL PROBLEMS

How far any ability shown by an individual is acquired from the environment and how far it is innate is notoriously difficult to assess. Even in the case of general intelligence where the matter has been investigated on a considerable scale over a long period, wide divergences of opinion still exist, in spite of general agreement that intelligence is the product of *both* heredity and environment.

The first systematic attempt to study heredity was made by Francis Galton, the pioneer of scientific psychology in England. Himself related both to the Darwin and the Huxley families, he concentrated his interest, after some years as a doctor, biologist and geographer, on the problems of heredity. He collected data on 997 eminent men in 300 families and demonstrated that the number of eminent relatives was far greater than would be expected by chance. He concluded from this that genius was inherited. When the first edition of his *Hereditary Genius* (1869) was published, the idea of the hereditary transmission of human powers was unfamiliar. As he wrote in the preface to the second edition which appeared in 1892, 'the human mind was popularly thought . . . to be capable of almost any achievement if compelled to exert itself by a will that had power of imitation.' Failure to learn was liable to be attributed to a lack of diligence on the part of the pupil or to incompetence on the part of the teacher. Only 60–70 years ago the Board of Education regarded backwardness in school subjects as the teacher's fault; for, if the child was below the level considered normal for his age the school did not receive its grant for that child. The practical effect of such a policy was that the dull children were often overdriven and the gifted neglected.

But Galton's belief that the fact that talent ran in families necessarily proved that it was inherited, was also open to criticism. The abilities shown by eminent men might, his critics argued, just as easily be explained by the intellectually rich and stimulating environment which many had enjoyed. Unfortunately for the study of the inheritance of human abilities, the closer the genetic relationship, the greater the chance of the members of the family sharing the same home and social environment for long periods, and it is therefore impossible to determine how much of the family resemblance is the result of common environment and how much is the result of heredity. In fact, wide differences between individuals in the same family may be a more potent argument for heredity than the moderate degree of resemblance which is usually found. The environments of two brothers, for example, are never exactly the same. However, environmental differences can hardly explain the professor's son who is very dull, nor why one brother should be conspicuously bright and another average or below average.

A really scientific study of human heredity would require that either the heredity or the environment was held constant while the other varied. Nature has provided some help in holding heredity constant by producing identical twins. Since these arise from a single fertilised ovum, they have an identical set of genes, while fraternal twins spring from separate ova which happen to be fertilised at the same time and are genetically no more alike than ordinary siblings (brothers or sisters). Since fraternal twins share the same pre-natal environment and are born at the same time, they tend to be more closely associated than siblings. On the other hand, identical twins may acquire differences between conception and birth, e.g. the supply of maternal blood to each may not be equal. However, from a comparison of the average resemblance between identical and fraternal twins, one can at least tell whether or not heredity is of importance for a particular characteristic, and can sometimes estimate how important it is compared with environmental conditions.

Normally, identical twins are brought up in the same home and are particularly liable to be treated alike by their parents and friends; they tend to go around together and may be thought to influence one another in various ways. Occasionally identical twins are adopted into different homes and localities. The most extensive study so far reported was the investigation carried out by James Shields of the Medical Research Council who, with the help of the BBC, succeeded in locating 44 pairs. His research reported in a book published in 1962 convinced Shields of the importance of heredity in both intellectual and personality factors.

GENETIC STUDIES OF MUSICAL ABILITY

As in the case of intelligence, the heated arguments in the nature-nurture controversy have tended to give way to an appreciation that heredity and environment interact in the development and manifestation of musical ability.

Most psychologists of music would nowadays probably agree with Farnsworth (1958, p. 184) when he says 'It is now clear that neither nature nor nurture can alone make a musician. Both must be present before musical and other abilities can emerge.' Nevertheless, opinions differ as to which side should be stressed. Wing and Drake emphasise the importance of innate factors (see Wing, 1963; 1948, p. 77; Drake, 1957, p. 13). On the other hand, Farnsworth himself and Lundin with his 'interbehaviourist' theories (see p. 173 below), seem to lose no opportunity of pointing out the contribution of environmental factors. This may be partly because they have in mind the need to qualify the dogmatic statements of Seashore, Schoen, and Kwalwasser on the hereditarian side. For example, according to Schoen (1940, pp. 161–3) 'Musical talent is first an inborn capacity. Artistic musical performance rests ultimately on innate, inborn equipment'; while Seashore (1919, p. 6) stated, 'Not only is the gift of music itself inborn, but it is inborn in specific types.'

It is of practical importance to try to reach some estimate of how much musical ability can be improved by a favourable environment and efficient teaching. Policy decisions have to be made on how far music lessons should be available to every child, however poor his ability. Since music is not a 'bread and butter' subject like reading or arithmetic, it could be argued that if musical talent is largely innate, it is not worth while schools spending too much time on the ungifted. It would be better to encourage the unmusical to pursue other more profitable activities, allowing teachers to devote more time and effort to discovering and fostering the talent of the gifted. Again the teacher may feel reassured if the failure of some of his pupils to progress can truthfully be ascribed to innate lack of musical ability. The parent may wish to know whether it is worth spending time and money on private music lessons for a child who has not shown definite promise of ability.

The influence of the environment on a relatively specialised ability, such as musical ability, might appear somewhat easier to determine than on general intellectual ability. Indeed, before the days of the record player and the radio, it might have been comparatively simple to obtain an accurate estimate of the music a child heard at home and of the training he had received from school or individual lessons, though interest in music would still have been an important factor. The greatly increased opportunities of hearing good music professionally played in the home which nowadays exist, have enriched the environment to an extent that is difficult to assess.

Two approaches to the study of the heritability of musical ability have been made: (1) genetic studies of family resemblances, and (2) attempts to assess how far environmental influences affect the performance of tests or other musical tasks.

In the next chapter investigations of family resemblances will be summarised. Chapters XIII and XIV review attempts to study the biological mechanisms by which musical capacity could be passed on from parent to child and racial differences. Chapters XV to XVII deal with various types of environmental factors.

XII

Genetic Studies of Musical Ability

The earlier attempts to investigate how far musical ability is inherited were mostly pedigree or questionnaire studies. Since the development of musical ability tests some research has been carried out in which parents and children or siblings or twins have been tested and their scores compared.

PEDIGREE STUDIES

Among the eminent and illustrious men studied by Galton (1869) were 120 musicians. By an 'eminent' man, he meant 'one who had achieved a position that is attained by only one person in 4,000'. By 'illustrious' he meant 'men whom the whole intelligent part of the nation mourns when they die; who have, or deserve to have, a public funeral; and who rank in future ages as historical characters. Such men occur only once in a million or once in many millions.' Galton considered only seven of his musicians were 'illustrious': Bach, Beethoven, Haydn, Handel, Mendelssohn, Mozart and Spohr. Few people today would give Spohr a place on such a restricted list. The claims of Mendelssohn over those of Schubert, for example, might also be disputed. Galton used a list drawn up by 'a literary and artistic friend' as being the most original and eminent musicians upon record, having found it impossible to obtain a list of first-class musicians that commanded general approval, of a length sufficient to show the influence of heredity.

26 out of the 120 or about 1 in 5, had had eminent kinsmen. As Galton included nine members of the Bach family and two members of four other families, the 26 belonged to only 14 families. In the case of Mendelssohn and Meyerbeer the eminent relatives had achieved their success in careers other than that of music.

A more extensive genealogical study confined to musicians was carried out by Feis (1910). He attempted to collect information about the parents and children of 285 famous musicians, but found the data on the maternal lines very hard to obtain, so that the material he could assemble was seriously incomplete. Moreover, many of the musicians did not produce children.

An investigation of contemporary musicians and students at the Juilliard Graduate School of Music, was carried out by Amram Scheinfeld for the first edition of *You and Heredity*. His virtuosi group in the following table included such outstanding performers as Yehudi Menuhin and Artur Rubinstein.

TABLE XII.1 *Approximate percentage of relatives with some degree of talent*

<table>
<tr><th rowspan="2">Group</th><th rowspan="2">No. in group</th><th colspan="3">Percentage for</th></tr>
<tr><th>Fathers</th><th>Mothers</th><th>Siblings</th></tr>
<tr><td>Virtuosi</td><td>37</td><td>75</td><td>50</td><td>50</td></tr>
<tr><td>Opera singers</td><td>36</td><td colspan="2">67</td><td>40</td></tr>
<tr><td>Students of music</td><td>50</td><td>58</td><td>74</td><td>70</td></tr>
</table>

An analysis of the incidence of talent in the three groups showed that where both parents had musical talent, more than 70% of the brothers and sisters (in addition to the individual reporting) also had talent. Where only one parent was talented, there was talent in 60% of the siblings. Where neither parent was talented, only 15% of the brothers and sisters had talent.

Among the virtuosi instrumentalists, the majority had talented parents one or both. Yet quite a number reported no talent in either parent. Nor did the differences in the family backgrounds, or in there being both parents, one parent, or neither parent talented, seem to have anything to do with the calibre or quality of musicianship shown by the individual.

Sergeant (1967) has also collected some data on the percentages of his professional music groups and of his control group (see p. 72) who had parents that could play an instrument:

	Music groups	General students
At least one parent able to play	Approx. 80%	62%
Both parents able to play	30–42%	24%
At least one parent a professional musician	20–45%	6%

QUESTIONNAIRE STUDIES

In the 1920s several genealogical studies were carried out on the Continent in which musical ability was assessed entirely or partly by questionnaires.

Haecker and Ziehen (see Revesz, 1953) found that the chance that a child will be very musical is 86% where both parents are talented, about 60% when one parent is musical, and about 25% when both parents are unmusical. Remarkably similar percentages were obtained by Heymans and Wiersma (see Revesz, 1953). That no less than 25% of the children of unmusical parents are described as being very musical might be partly due to the parents making more generous estimates of their offsprings' ability than of their own. Also, the ability may have been inherited from grandparents.

In order to give some degree of objectivity to the assessment of his subjects' musical ability, Jan Mjoen of Oslo (1926; 1934) used a musical index graded from 0 to 10. In the following table those described as (P) 'Poor' were rated between 0 and 2. Their ability was limited to being able to recognise a tune (2). The (M) 'Musical' (3–7) ranged from those who knew when they sang or played out of tune (3), through (5), holding a second part, to (7) being able to improvise a second part. The (S) 'Superior' group (8–10) were, at the least, able to play by ear, while the most talented of all could compose and play several instruments.

TABLE XII.2

Parents	Number of parents	Number of children	%age of children S	M	P
S × S	7	23	72	28	0
S × M	40	175	60	34	6
S × P	9	34	26	37	37
M × M	30	113	39	49	12
M × P	21	75	7	40	53
P × P	7	22	0	10	90
Total	114	442			

It is apparent from this table that the higher the average talent of the parents, the higher the average talent of the children is likely to be. Mjoen believed that this could be regarded as a demonstration of the inheritance of musical ability. The environmentalist might object that the musical environment provided by the parents is likely to vary roughly with their own talent or lack of talent. Mjoen, however, also presented evidence that where both parents are musical (grade 5 or above), the proportion of children who are musical corresponds to the number of grandparents with talent. Thus, where three of the grandparents are musical, 90% of the children are likely to have musical aptitude, if only one grandparent has talent, only 50% of the children may be musical. Mjoen, therefore, concluded that it is the quality of the stock rather than the quality of the parents which determined the ability of the children. He quoted the following case as an illustration. The four children of a grade 4 parent married to a grade 5 were rated as grade 5. One married a woman with little musical aptitude (grade 2); their children averaged grade 3 in musical ability. One daughter's spouse was very unmusical. One of their children was rated grade 2 and a second, grade 4. On the other hand, a daughter who married a man with grade 6 talent and siblings whose ability averaged 7 produced seven children whose capacities ranged from 5 to 10. Her sister also married a man whose aptitude was rated as 6. However, his siblings were less musical (average 3). The ratings of their three children were only 3, 4 and 5.

Among the family trees studied by Mjoen was that of the Norwegian composer, Halfdan Cleve. His father, who was very musical and came from a musical family, married twice. His first wife was unmusical and came from unmusical stock. None of their five children showed musical aptitude. His second wife, however, was musical. All their five children were above average musically, and one, Halfdan Cleve himself, very gifted. Halfdan Cleve married a well-known pianist, who came from a musical family, one of her siblings being highly talented. The four children resulting from their marriage showed promise of exceptional talent.

One performer of international repute conspicuously failed to conform to Mjoen's hereditarian theories. Her family was evidently quite without talent. However, Mjoen later discovered that she was the illegitimate daughter of an eminent musician and had been adopted by unmusical parents.

Questionnaire surveys however carefully carried out, depend on the accuracy of the replies. Where instrumental playing is taken into account, less than justice may be done to individuals who have lacked the opportunity to learn to play.

STUDIES BASED ON TESTING

Parent–Child and Sibling Studies

As might be expected the first tests to be applied in genetic studies of musical ability were Seashore's.

In fact, during the years of experimentation which preceded publication of the original Measures, one of Seashore's students, Felix O. Smith (1914), applied an early form of the pitch tests using tuning forks to groups of schoolchildren. The results were disconcerting to Seashore, who was a strong believer in the inheritance of musical capacities. The correlation obtained from the scores of siblings was ·43 for those without practice and ·48 for those who had been given some practice. This figure was indicative of a moderate degree of connection between siblings on pitch discrimination. However, when he compared the younger children with unrelated children of the same age and sex as the elder of his pairs of siblings, he obtained a correlation of ·53. Near zero correlations are to be expected from unrelated children, and this was found in the Syracuse sibling studies (see p. 121).

Shortly after the publication of the Seashore measures, Stanton undertook a pioneer genetic study to explore the possibilities of using standardised tests for such a purpose. It seemed to her advisable to begin with the families of six well-known American musicians, as they were more likely to be co-operative than unselected persons. The basis of her sampling was thus rather narrow, though some members of the families were in fact unmusical. She interviewed and tested 85 persons, ranging in age from 8 to 80. In five of the families she was able to study three generations. She used the Seashore records to assess intensity and time discrimination and tonal memory and, by way of variety, tuning forks to test pitch.

Stanton published her results in a monograph (1922). She did not attempt to correlate the parents' and children's scores or those of the siblings. Instead, talent profiles were drawn up for each family. She herself raised the question of how far the percentile ranks of her older and younger subjects were strictly comparable, and thought that further research might show a need for different norms for subjects, aged 45 to 65, and for those over 65. Stanton's graph plotting pitch scores against age shows a fall in the 40s, a recovery around 50 and a marked decline after 60. Tonal memory showed little effect of change with ageing. It would have been interesting to have had correlations at least for this test (the one most likely to be valid in the sense of relevant to musical progress) and for the total Seashore scores with the effect of age eliminated by statistical means.

Stanton presented some interesting supplementary data in which the overall Seashore talent profiles are tabulated against various environmental factors. These will be considered in Chapters XV and XVII. She also examined the percentile ranks on each of the Seashore subtests of the children produced by the mating of parents of various levels of talent. Most of the parents were above average, but their children tended to be superior to the average of the parents. For example, on tonal memory, of seven parents one of whom had superior talent and one average, fourteen of the children were superior, two average and one poor. This may have been partly due to the children being more adaptable to the test situation. With a more heterogeneous sample, more interesting results might have been obtained. Stanton contented herself with the cautious conclusion that it was 'not impossible' that the inheritance of the capacities measured by the Seashore tests followed Mendelian principles.

Ruth Friend (1939) made an interesting, if optimistic, attempt to apply the Seashore tests of pitch, intensity and consonance as adapted by McGinnis (1928) individually to 20 boys and 22 girls at a kindergarten school. Their average chronological age was 5:3 and their mental age averaged 5:11. The tests were simplified by shortening the amount played and by increasing the interval between judgments. The children were given the tests twice within a week. The coefficients of reliability on retest were: pitch ·57, intensity ·61, consonance ·51 and combined tests ·78. However, correlations between the test scores and ratings by three teachers of the children's musical ability was low (·15). All three teachers had good opportunities of knowing the children and their ratings agreed quite closely. The sum of the parents' ratings of their children's ability also showed a low correlation (·26) with the children's Seashore scores. The parents were tested with the three Seashore tests, applied once. The following correlations were obtained between the mean of the parents' scores and the child's mean for two trials: pitch ·14, intensity ·46, consonance ·11. The correlations with the 25 fathers were: pitch ·02, intensity ·16 and consonance ·04, while the 35 mothers' scores correlated with the child's ·09 for pitch, ·28 for intensity and ·08 for consonance. The highest correlations were thus obtained for intensity, the most reliable of the single tests as used in Friend's study. This might be due to the children having a clearer conception of 'loud' and 'soft' than of 'high' and 'low'. Seashore himself (Hattwick and Williams, 1935) came to the conclusion that even children of between 6 and 9 had not adequate concepts of 'high' and 'low', even after practice and exploration, to undertake the Seashore pitch test. Friend herself doubted if the answers to the consonance test (a test no longer part of the Seashore measures) were based on any real aesthetic judgments. Low negative correlations were obtained between the parent's estimate of the child's musical environment and his Seashore scores. Given satisfactory tests of musical aptitude of young children, an investigation such as Friend's would be of considerable interest.

A more recent investigation using a modified version of the Seashore tests was carried out by Professor Luigi Gedda (1961) of the Mendel Institute in Rome. He and his colleagues tested a group of choirboys chosen from the Sistine Chapel and as many of their relatives as possible. The boys were between 11 and 15 years of age. A group of boys of the same age, but not musically selected, acted as controls. The choirboys were found to be clearly superior in their discrimination of pitch, intensity and time to the control group. Their relatives were also superior. With the tonal memory test, Gedda asked his subjects merely whether the second playing was the same or different from the first. 80% to 90% of the choirboys' answers were correct. But both their relatives and the controls had definitely fewer correct answers.

UNIVERSITY OF SYRACUSE STUDIES

Kwalwasser (1955) summarised the results of three student researches based on the use of his tests. The correlation of the scores of 255 pairs of siblings was ·48. The 71 pairs of brothers' scores correlated ·56, while the 65 pairs of sisters' scores correlated only ·46. With 151 Negro siblings, a correlation of ·53 was obtained. In both cases, when random pairings ignoring the blood relationship were made, only zero or insignificant correlations were found.

Kwalwasser also mentions three attempts to compare parents and children which were made by experienced and well-liked music teachers. Unfortunately, all three found the mothers, though interested, had difficulty in arranging time to be tested, while many of the fathers seemed disinterested and uncooperative. Kwalwasser does not give the correlation coefficients, but they were 'lower' than with the siblings. This may have been due, he suggested, to the unfavourable attitude of the parents, to the considerable difference in age, or to a deterioration of hearing on the part of the parents.

SHUTER'S STUDY

The tests used in previous investigations were not the most satisfactory from the point of view of validity and reliability. Shuter, therefore, used the well-validated Wing tests for a study of parents and children, and of twins, carried out at the University of London Institute of Education.

The aim of the first part of her investigation was to compare the musical ability of parents and children, to see what connections existed which might be related to hereditary and/or environmental factors. Information on the latter was obtained by questionnaires (Shuter, 1964; 1966).

The main group tested consisted of 54 pupils of a mixed grammar school and 63 of their parents. Thirteen pupils of a grammar school for girls and 15 of their parents were also tested. Both schools were situated in the Home Counties. The children's ages ranged from 11:1 to 18:1 (one elder sister, aged 21, was also included). To enable the children's Wing scores to be compared with those of their parents, the raw scores were converted into Musical Quotients (MQs) by the formula given in the Wing Test Manual. As Shuter also wished to consider the first three tests separately from the four appreciation tests, corrections for age were made in accordance with the differences which appeared when the raw scores were plotted against age.

The desirability of including parents and children at all levels of musical ability was stressed to the schools beforehand. Unfortunately, a volunteer group tends to become a self-selective one, since adults are more likely to attend if they themselves are musical, or have some interest in music or feel that their children may show talent. In fact the average MQ of the parents was 112 and that of the children, 132.

As is well known, any group so 'selected' from the top end of a scale is very likely to give depressed correlations. However, when the MQs of both parents were averaged and compared with their children's, a correlation of ·475 was obtained. With the 14 pairs of siblings, the correlation was ·496. These figures were close to the ·5 'typically' reported between parents and child and between siblings on intelligence. However, comparing each child with each parent (number of cases = 100) yielded a correlation of only ·290. But this rose to ·364 when the results from the girls' school, where the self-selective effects were particularly marked, were omitted. Shuter therefore considered that with a more representative sample of parents and children the level of correlation would rise rather than fall. Wing tests 1–3 correlated ·157 – even lower than appreciation tests considered separately (·260).

Shuter also compared the MQs of the fathers with those of their wives in the 25 cases where she had been able to test both parents. The figure obtained was quite low: ·331 compared with the range of correlations, from ·3 to ·7, mentioned by Cattell (1950) as found with married couples for a variety of traits such as height, intelligence and interests. One might have supposed that amateur musicians who devote a high proportion of their spare time to music would stand quite a good chance of meeting and marrying women with similar interests.

TWIN STUDIES

Kwalwasser (1955) mentions a twin study carried out in which a correlation of ·77 was obtained on the K–D tests and of ·84 on the Kwalwasser–Ruch Test of Music Accomplishment. Achievement tests often do correlate more closely than for example intelligence tests, especially among fraternal twins. Kwalwasser does not say how many of the 25 pairs of twins were considered identical.

The validity of studies comparing identical with fraternal twins obviously depends on the accuracy with which the two types of twin can be distinguished. Investigators who have been able to supplement their own observations of the appearance of the twins by such objective measures as blood grouping and examination of finger prints claim to be able to identify as fraternal or identical over 97% of twins. One such extensive investigation was the Michigan Twin Study (Vandenberg, 1962) for which batteries of cognitive and perceptual tests were administered to some 33 pairs of identical and 43 pairs of fraternal twins. Vandenberg included Seashore's pitch, loudness and rhythm tests and Wing's pitch and memory tests. Only in the case of the rhythm and memory tests were there significant differences between the scores of the two types of twin. The 'heritability indices' intended to show the percentage by which heredity contributed to each variable were as follows:

Seashore	Pitch	00%
	Loudness	44%
	Rhythm	52%
Wing	Pitch	12%
	Memory	42%

It is surprising that the two pitch tests should apparently be so little subject to hereditary control. Acuity of hearing, which is, however, not closely related to pitch discrimination (see p. 197) proved highly heritable, but only in the case of the right ear. Vandenberg himself suggested that it might only be the exceptional talent of the great composers that has an hereditary factor.

Stafford (1965) has re-analysed some twin data collected by Thurstone based on testing 48 pairs of identical twins and 54 pairs of fraternal twins. In his re-analysis the contrast between the types of twin was found to be significant for the Seashore pitch tests. In this case the memory test was found to be less hereditable. Both Thurstone and Stafford found significant degrees of contrast in the case of the Seashore Rhythm Test.

In the second part of Shuter's investigation, 50 pairs of twins ranging in age from 9 to 16 were tested. They were classified as identical or fraternal by general impression and careful inspection. In addition, Shuter obtained Wing scores for 11 pairs of young adult twins who were being studied by the Medical Research Council Psychiatric Genetics Research Unit for another purpose.

Many of the twins were well below average in musical ability and found the appreciation tests particularly difficult. Shuter therefore considered that MQs based on the first three Wing tests to be more reliable. As would be expected if heredity is a factor in musical aptitude, the average difference in MQ between each individual and his co-twin was lower among the identicals and a higher proportion of the identicals had lower intra-pair differences. In the following tables each class of

TABLE XII.3 *Distribution of intra-pair differences*

	Musical quotients								
MQ	Identical				Fraternal				
points	Children		Adults		Children			Adults	
difference	Boys	Girls	Men	Women	Boys	Girls	Mixed	Men	Women
0–4	4	2	2		1	3	1	1	
5–9	1	2		1		2	1		
10–14	2	2	2		3	2	2		
15–19		1	1	1	2	1	4		1
20–24	3	2		1	1				
25–29					2	2	1	1	
30–34		1				1			
35–39									
40–44									
Over 45						1			
Totals	10	10	5	3	9	12	9	2	1

twin is subdivided by sex and by age, though the number in each category is quite small.

Shuter in her thesis (1964) compared the extent of intra-pair differences with the twins' attitude to music, amount of training and of listening. No consistent relationships emerged. The majority of the subjects had never had music lessons and only very few ever attended concerts or listened to classical music in their homes. Among the identical twins, if one had received music lessons, so had the other. A difference of 22 MQ points was found among a highly talented pair of boy twins, judged to be identical. They had both been studying the piano for 4 or 5 years and did a good deal of practice. On the other hand, the two pairs of identical girl twins with the most discordant scores (differences of 23 and 33) had never had music lessons. Among the fraternal twins, the larger differences were found among those who were having lessons and who played for about equal periods, as opposed to those who were similar in not playing at all. But the numbers were too small to be conclusive as to whether or not the training might have caused, or increased the differences.

TABLE XII.4 *Means and medians of intra-pair differences*

	Musical quotients							
	Identical				Fraternal			
MQ	Children		Adults		Children			
points	Boys	Girls	Men	Women	Boys	Girls	Mixed	Adults
Means	10·90	13·10	9·20	15·67	16·33	20·04*	13·67	15·67
Medians	9·00	12·00	11·00	15·00	15·00	12·50	15·00	18·00

* Would fall to 13·86 if the pair with the difference of 88 points were excluded.

When the MQ of each identical twin was compared with his co-twin's, the correlation was ·84 for the 20 child pairs and ·79 when the 8 adult pairs were included. With the 20 pairs of fraternal twins of the same sex the correlation was ·72. Little difference was made to this figure by adding the results from the 9 pairs of boy-girl twins and the 3 adult pairs. The heritability indices showed that in the case of the children heredity contributed 42% to their musical ability and in the case of all the subjects, 26%. The contrast between the two types of twins was considerably less than has been reported on intelligence tests, where correlations of over ·9 have been obtained for identical, and of ·5 or ·6 for fraternal, pairs. But factors such as the level of difficulty of the tests may have affected Shuter's results.

In spite of misgivings about the reliability of the Wing subtests with such unmusical children, Shuter correlated the pitch and memory tests separately to see whether or not the results would confirm Vandenberg's. On the pitch test, the intra-pair correlation of the fraternals was higher than that of the identicals. But on the memory test the identical twins correlated ·77 and the fraternals ·50. The heritability index showed a contribution by heredity of 53%. (This rose by 64%, appreciably higher than Vandenberg's figure, when it was calculated by the formula used by him.)

Since the Wing pitch test has been found to correlate quite highly (see p. 198) with the memory test, it is surprising that the two tests apparently differ appreciably in the extent to which they are under hereditary control. As some measure of pitch discrimination would appear to be a pre-condition of melodic memory, one might have expected that, if any difference was found, the pitch test would show the higher degree of heritability.

IDENTICAL TWINS BROUGHT UP APART

Only one of the earlier studies of identical twins brought up apart included results of testing musical ability. This was concerned with a male pair brought up in Glasgow to whom Vernon administered the K–D tests. One twin apparently showed a definite aptitude for music. At the age of nine he had chosen to study the violin and had made satisfactory progress. After five years however, he had to give up his music lessons due to his father losing his job. The other twin's only special talent was playing the trumpet, which he took up at 14. He played in small amateur dance bands. His K–D scores were much inferior to his brother's, which was considered 'not surprising' in view of the latter's superior musical education. 'However, there was no general resemblance between the patterns of their scores on the individual tests, such as would indicate a genetic basis to their musical talents' (Yates and Brash, 1941).

Shuter was able to test 5 of the pairs brought up apart who had earlier taken part in the study carried out by the MRC Genetics Research Unit. In describing her results, she adopted the pseudonyms used by Shields in his book (1962) describing the MRC investigation.

Shuter found the following differences in Wing scores between each twin and his co-twin: 2, 2, 12, 15 and 20.

One of the pairs whose total Wing scores differed by only 2 points had a rather similar musical background. Aged 42 when tested, Jenny had been brought up in a London suburb and Kathleen in a small seaside town, where she was regularly visited by Jenny. Both claimed to be interested in music; both had had one or two years of piano lessons during childhood, though Kathleen claimed to listen to as much as six hours of classical music a week.

Both of the other pair of twins with a difference of only 2 points in their Wing scores had grown up in families where other members of the family played in brass bands. They had been brought up in a large town in the North of England and were aged 38 when tested. But while Francis had actually played the cornet and other brass instruments in a band for 24 years, Foster had not had the opportunity of learning to play till only two months before he was tested. When his 9-year-old son had begun to learn the cornet, he had taken up the E-flat bass, so that they could help each other in their lessons. He seemed to be thoroughly enjoying playing. The twins' subtest results were very similar.

Test	1	2	3	4	5	6	7	Total
Foster	11	18	16	7	5	8	8	73
Francis	12	17	16	9	8	6	7	75

The difference of 12 was found with an older pair of twins (age at testing = 57) who had been brought up apart in mining villages in the North of England. Jim whose score was only 56, had moved South as a young man. He was more interested in music than his co-twin and his tastes were more sophisticated than Robert's. He attended concerts at the Royal Albert Hall and listened to classical music. He liked ballet music and Tchaikovsky, while Robert preferred Strauss.

Jacqueline, whose Wing score was 79 compared with her sister's 64, had enjoyed a much superior musical background as a child. Her adoptive father conducted a church choir till the age of 80. There was much music-making at home. Jacqueline herself had piano lessons for seven years from the age of 14 and later conducted a Townswomen's Guild choir. Beryl, on the other hand, had never had music lessons, though she belonged to a choir. No one in her home played.

The largest difference in Wing scores, 20 points, was found among two female twins tested when 43. Separated at birth, they did not meet from the age of 3 till they were 39. Their father played the piano and had gained RCM diplomas when young. Olive, who was brought up by her own parents, had piano lessons for two years from the age of 10. From the age of 14, she listened to music, mostly light classical or organ music, every evening. Madge was brought up by an uncle and aunt who encouraged her to become a piano teacher. She gained the ALCM and ATCL diplomas, but at time of testing had given up music teaching and had become, like her twin sister, an assistant nurse. It is possible that some of the higher Wing score made by Madge was due to her feeling more confident at being able to do the tests than was her sister. Both scored well above the average and some of Olive's scores on the subtests were very close to Madge's:

Test	1	2	3	4	5	6	7	Total
Olive	10	27	17	8	14	5	9	90
Madge	17	26	25	9	13	10	10	110

They had similar tastes in music, as in clothes and books, as Shields (1962) noted: 'One mentions Tchaikovsky's first, the other Rachmaninov's second piano concerto as among her favourite compositions. Olive likes the Messiah best of all. On top of a pile of music in Madge's room was a copy of the Messiah. Although she plays only a very little herself, Olive says "music is a big influence in my life".'

These studies do provide some evidence of genetic factors in musical ability, but so far they do not emerge as clearly as in the case of intelligence.

XIII

How is Musical Capacity Transmitted?

The studies reviewed in the last chapter were primarily concerned with the question 'How far is musical ability innate?' But we can also ask 'By what biological mechanisms is musical capacity transmitted?' The latter question assumes, of course, that musical ability is to some extent under genetic control.

While Galton was engaged on his study of the genealogies of eminent men, an Austrian monk was patiently and systematically investigating the effects of crossing different strains of garden peas. Pure strain peas with a pair of genes for tallness were, he found, always tall. When crossed with pure short peas, the hybrids were always tall, the tall gene being the effective or 'dominant' one. More excitingly, when two hybrids were crossed, he obtained one pure tall, two hybrids (tall in appearance, with one tall and one short gene) and one pure short one.

The genetic laws of dominant and recessive genes which Gregor Mendel discovered were, however, published only in an obscure journal and their true significance was not recognised till 1900.

It was very natural that when the work of Mendel became known, geneticists should have been eager to find out whether the genes governing the transmission of human characteristics could be identified. This was relatively easy in the case of the kinkiness of hair where the effects were clear cut. But the number of psychological traits traceable to single genes was, as Cattell complained in his book on Personality (1950), 'regrettably small'.

In most cases, complications of three sorts are found:

1. There may be two or more different genes which produce identical or seemingly identical effects.

2. The same gene may have different effects in different individuals or under different environmental conditions. This may be due to variations in the 'penetrance' and 'expressiveness' of a gene. Every gene has

a certain degree of penetrance, and for a given degree of penetrance, may or may not be expressive in the individual.

3. Moderator genes may determine when the activity of a particular chain of genes will begin or end.

Characteristics dependent on dominant genes will run in families recurring generation after generation, since only one parent is needed to pass them on. Although the degree to which they are shown in any given individual may vary, an examination of family trees over several generations is likely to reveal inheritance of a dominant gene. Traits determined by recessive genes cannot be identified by following them through several generations since they usually occur suddenly in a ineage as the result of the mating of parents of similar genes. The parents and the children of the individual affected rarely show the recessive trait but it also appears in about one in four of the affected person's siblings.

Special patterns of sex-linked inheritance are sometimes found due to the two 'sex' chromosomes. Besides 22 pairs of matching autosomes, human females have two smaller X chromosomes and males one X and one very small Y chromosome. Because the X chromosome is much larger than the Y chromosome, for many genes on the X there are no corresponding genes on the Y. A woman with a recessive gene on one X chromosome only will not be affected by it, but a man with such a gene on his X chromosome will have no normal alternative gene on his Y chromosome and will be affected by it.

The man will not transmit the trait to his sons, but his daughters will all be carriers since they have received an X chromosome from their father. When they come to have children, half their sons will inherit the trait and half their daughters will be carriers.

If a gene were linked with the Y chromosome, all a man's male descendants would be affected, but none of the female line. No certain cases of Y-linked inheritance are yet known, though some will almost certainly be found (Carter, 1962).

In sex-limited inheritance, men only show the trait (e.g. a certain form of premature baldness). Women can inherit it from their fathers and transmit to their sons, but do not show it themselves, perhaps because the glandular makeup of the two sexes may govern the way in which the gene expresses itself. The distinction from sex-linkage is readily made by the fact that in sex-limitation men can transmit the condition to their sons.

THE BIOLOGICAL MECHANISMS OF THE INHERITANCE OF MUSICAL CAPACITY

The first attempt to apply Mendel's principles to musical capacity seems to have been made by C. C. Hurst (1912). He observed the population of an industrial village in Leicestershire over a number of years, noting how many of the children took after their parents in 'musical temperament' as well as a number of physical characteristics such as hair colour, eye colour and left-handedness.

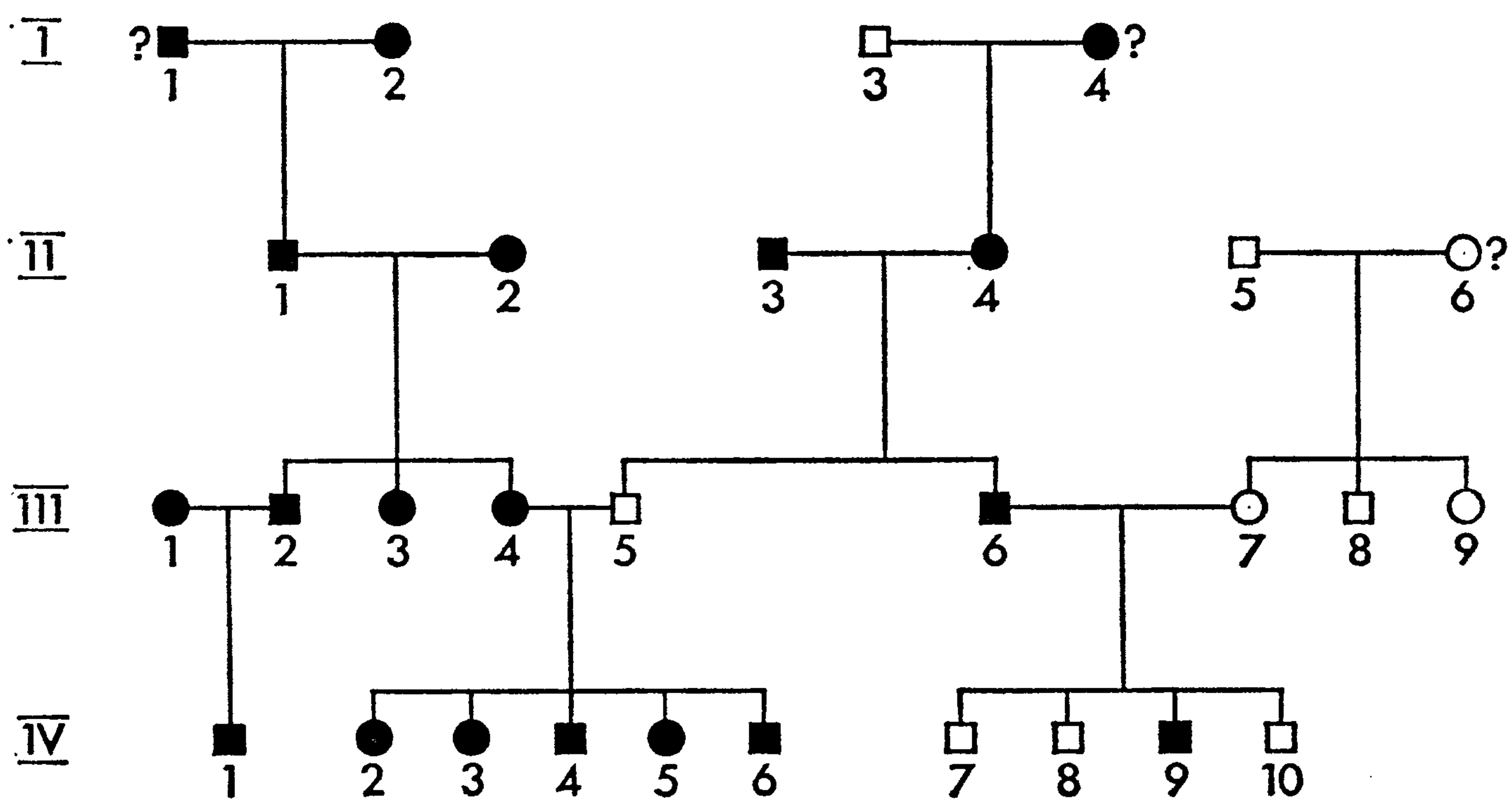
By musical temperament he meant ability to be able to pick up a tune after only a few hearings, spontaneous singing early in life and a 'natural feeling for harmony'. According to Hurst, when both parents are musical, all the children are musical. When neither parent is musical, signs of aptitude will be shown by none or by only a few of the children. When only one parent is musical, either none of the children has ability or 50% are musical and 50% unmusical. Hurst concluded that musical ability was a recessive trait, lack of aptitude might be due to an inhibitory factor preventing the expression of the musical talent which is hypostatically present in everyone. Do the tone deaf, he wondered, suffer from a double dose of the inhibitory factor?

Drinkwater (1916) tried to apply Hurst's conclusions to the pedigree of a family that had produced several distinguished organists. When both parents were musical, all the children – no less than 10 in one case – were found to be talented. When one parent was unmusical but of musical stock, one half of the children were musical. But in the case of this family and of others mentioned by Drinkwater, when only one parent was unmusical, the proportion of children showing talent often exceeded theoretical expectation. Drinkwater supposed that the mothers (the parents classified as unmusical) might, in fact, have had some ability, but that it had been eclipsed by their husbands' high degree of professional talent. The lack of a graded method of assessing musical ability leaves such questions unanswered. But the children of professional musicians would have excellent opportunities of learning, and some with only moderate gifts might have become competent musicians.

Two pedigree studies relating to musical ability were included in a series published in the United States.

Northrup (1931) studied three generations of one family. In the first generation the father was highly talented and the mother unmusical. The second generation consisted of five musical children and one who was 'musically inclined'. A musical son married an unmusical woman and produced three definitely musical children and one moderately musical one. The other musical son married a fairly musical woman; their child was musical. A musical daughter married a man without talent; their four daughters showed ability but their two sons did not.

Reser (1935) endeavoured to follow musical ability through five generations of one family. Unfortunately her main criterion of musical ability seemed to be the possession of a fine singing voice. It may have been the voice and not the ability to use the voice which was being handed down. For the first two generations of the family she was dependent on descriptions by living members of generations III to V. These may not have been accurate. The musical trait appeared to be passed on through the generations without a break; it was thus apparently inherited as a Mendelian dominant trait, but the progeny of the



Squares represent males, circles females.
Blank □ and ○ represent no ability in music.
(Adapted from Ashman, 1952)

non-musical members of the family were not given and data about them would have been equally valuable.

An interesting attempt was made by Ashman (1952) to trace the biological mechanism of ability to carry a tune, through four generations involving members of three families. Though Ashman quotes Seashore scores for III 6 (PR = 51), III 5 (PR = 15 to 20) and for III 3 (PR = 76) (see pedigree chart above) most of his information appears to be based on his personal knowledge of the families. His report includes some comment on the individuals concerned.

The brothers, 5 and 6 from generation III, lived up to the ages of 8 and 10 respectively in the mountains of Virginia. They had no schooling and no contact with music. As soon as they were sent to school, 6 rapidly learned to reproduce simple melodies from memory, while 5 never learned to sing more than an occasional two or three bars in tune.

In generation IV, male 7 enjoyed music at school and listening to gramophone records at home. He went to a college where behaviourist ideas prevailed and where musical ability was regarded 'as an outstanding example' of an acquired ability. He practised hard, but wholly without success.

IV 10 was encouraged by his mother, though she was apparently not herself musical, to take lessons. Though he made rapid technical progress, he had no ear for music and no musical memory. His brother, IV 9, received no encouragement to learn and no musical training, yet he could recall simple melodies.

Ashman put forward the following biological explanation of the pedigree.

Individuals known to have had superior musical ability, e.g. II 1 and 2 and III 3 and 4 have received from both parents an incompletely dominant gene.

Those in whom musical memory is apparently either absent or very deficient have a pair of recessive genes (as Hurst had envisaged). He concluded that 'simple' memory for melody is possibly determined by a multi-factor gene, having other effects which are of survival value. (This is reminiscent of Darwin's idea that music originated as a means of attracting a desirable mate.) Melodic memory is, however, probably not such a simple factor as Ashman supposed, but at least the evidence cited elsewhere in this book shows that it is likely to be closely related to musical ability.

Ashman was, of course, right in thinking that it is likely to be easier to trace the genetic mechanism of a distinctive characteristic that can be isolated and identified as present or not present. Tone deafness seemed a promising condition for a genetic investigation. Since an inability to perceive or reproduce pitch patterns in a normal fashion is of concern to teachers of speech as well as to music teachers, a joint investigation of tone deafness was carried out by Dr H. Kalmus of the Eugenics Department and Professor Dennis Fry of the Phonetics Department of University College, London. A test was developed which appeared to discriminate efficiently between the tone deaf and the normal, at least among intelligent adults and adolescents. This was a distorted tune test consisting of the first two or more phrases of 25 well-known tunes. In one version the tune was played correctly and in the second the melody was distorted by the insertion of several blatantly wrong notes, the rhythm and tempo remaining unchanged. The subjects were asked to decide whether the tunes were played rightly or wrongly. Seashore's memory test, a number memory test and a pitch discrimination test were also used. Musical subjects occasionally marked a right tune as wrong because of some slight distortion in the recording or irregularity in the playing. This was an 'A' type of error. Consistent marking of wrong tunes as right seemed indicative of some deficiency in the subject. This 'B' type of error was given three times the weight of A and the scores were calculated as 3B–A (where B is the number of wrong tunes judged right and A the number of right marked wrong). The results showed a clear-cut division into two groups. 95% of some 1,200 subjects with marks of less than 14 were considered normal, and the other 5% tone deaf (Fry, 1948). Performance of the Seashore memory test was invariably found to be bad among the tone deaf. A significant but weaker correlation with pitch discrimination was found. No very strong degree of association was found with a number memory test. That the results of the distorted tune test were not unduly influenced by opportunity to learn the specific songs chosen was indicated by the fact that Continental students who did not know them made hardly more type B errors than the English.

The tone deafness (or tune deafness as Kalmus called it) appeared frequently to segregate in families and siblings in ratios indicating that it might possibly be caused by a single gene, possibly a dominant (Kalmus, 1949). However, in an article in *Scientific American* in 1952 he stated 'We do not know for certain whether tune deafness is caused by a single gene difference or controlled by many genetic factors.' He thought that there might be several types of tune deafness and that it was by no means independent of upbringing.

As a result of his enquiry into the musical aptitude of the relatives of professional musicians (see p. 116), Scheinfeld put forward a theory on the inheritance of conspicuous talent. He agreed with Seashore that musical aptitude has specific components such as the sense of pitch, or rhythm, etc., and that these have a constitutional basis. But while a foundation of such aptitudes is required before any talent can develop, something more is needed for real talent and virtuosity. Scheinfeld, therefore, supposed that the highly talented must have in addition 'certain rarer "special" genes which act either to intensify the effects of the more ordinary "aptitude" genes or to produce some unusual supplementary effects'.

In each of the three groups he studied the incidence of talent followed the pattern quoted on page 116 above. It is clear, stated Scheinfeld, that no single dominant gene, or no two recessive genes, could account for these ratios. A multiple-gene mechanism would be needed, and the simplest one which might fit the requirements would be, at the very least, two different dominant genes, passed on by only a single parent to a given child, or each parent could give the child one of the required genes.

Scheinfeld then applied this theory to Toscanini's family. The lack of talent reported in Toscanini's ancestry could be due to the paternal side of the family handing down only one of the required special 'talent' genes, while on the maternal side the other had been carried – neither being effective by itself. Both genes were brought together through the mating of Toscanini's parents, when the chance of the combination appearing in one of their children was one in four or less. To explain why so little of Toscanini's talent had been passed on to his children or grandchildren, Scheinfeld pointed out that even if his wife (described as 'mildly musical') carried one or both of the required 'talent' genes, the odds might still be only about fifty–fifty of any given child receiving the required double combination. In fact, only one of Toscanini's three children appeared to have any talent at all and only the child of this daughter, out of the three Toscanini grandchildren, showed any talent for music.

Scheinfeld's assumption that the great musicians differ from those of lesser talent in possessing a special kind of talent for music for which a special type of gene is required, is open to question. At least up to the level of high professional talent, there would seem to be no fundamental difference in kind (Shuter, 1964). Scheinfeld admits that other types of qualities – tenacity of purpose, performing ability, etc., are necessary for success. A combination of high all-round abilities with specifically musical talent might account for genius without postulating a special kind of talent. If, however, such a special kind of endowment does exist, Scheinfeld's assumption of a double dominant gene might conceivably be correct. In any case he puts it forward merely as a possible theory.

It would seem unlikely that any sex-linked mechanism could be involved in the transmission of musical aptitude when boys and girls make approximately equal scores on the Wing and other musical ability tests (see Chapter VIII). Sex-linked inheritance involves a much higher proportion of males showing the inherited characteristic than females. Thus in the case of colour-blindness some 4% of males are afflicted but only ·4% of females. However, Haecker and Ziehen's data seem to suggest that musicality may be inherited to a greater degree from the father than from the mother. 22% of the 74 cases of 'musically productive' individuals seem to have inherited their ability from both parents. In 25% of the cases the talent appeared to have been transmitted by the father alone, and in only 12% by the mother alone (Haecker and Ziehen, 1922, cited Revesz, 1953, p. 191). As we saw on page 118 it is, however, difficult to assess how far experience of music affected their results. It is possible that the child's opportunity of learning music and becoming musically productive may have depended on whether or not his father, rather than his mother was musical, in these Continental families of half a century ago. Swift found that the K–D scores of brothers were more closely associated than those of sisters (see Kwalwasser, 1955).

Shuter, however, found a clear difference between the agreement of the Wing scores between father and child as compared with mother and child, and between male, as opposed to female, twins. The correlation of the children's MQs was quite high (·627) with those of their fathers, but quite low (·258) with those of their mothers. Shuter scrutinised the data from the parents' questionnaires on the amount they played, on their music lessons, and their listening habits. However, even for these children of grammar school age, it seemed to be the mother, rather than the father, who set the musical environment. With the 10 pairs of identical boy twins, a correlation of ·90 was obtained. When compared with the correlation of ·73 for the nine pairs of boy fraternals, a hereditability index of 62% was found. Here again, the environmental data did not provide any explanation of the sex difference.

As far as is known, sex-linked characteristics can be passed on only from father to son *or* from father to daughter, but not to both. Some selection effect might be involved in Shuter's results, for example men may be more inclined to volunteer to be tested if their children resemble them in being musical and unmusical. The number of male twins was quite small. However, Shuter thought that there might be some connection between her findings and the higher proportion of boys classified as tone-deaf. Bentley (1954 and 1957) carried out an extensive enquiry among teachers to find out how many children were considered 'monotones'. The percentage of boys was much higher at the age of 7 and failed to decrease with age as much as among girls. At 12, 7% of boys, but only 1% or 2% of girls, were categorised as 'monotones'. Bentley himself thought that the 4% of individuals who are still monotones at 12 may be the same 5% of the adult population of both sexes who are tune deaf according to Fry. Fry, however, did not give separate percentages for the two sexes. But, why some 'monotones' outgrow their deficiency before the age of 12 would require investigation.

XIV

Racial Differences

It has sometimes been suggested that certain races, such as the Negroes, the Jews or the Germans, have special endowments in music. Thus Moss (see Klineberg, 1935) expressed the view: 'With the Negro, rhythm seems to be an innate quality, and he can extract music from any sort of instrument.' As Klineberg himself points out, however, Negroes themselves often question statements of this sort. For many Negro communities music plays an important part in life, but this is by no means universal. In rural West Virginia where Klineberg lived for some time he found very little interest in music among the Negroes; such interest as existed was directed towards modern American songs. In the larger cities it seemed that music meant no more or no less to the Negro than to his white fellow citizen.

In the 1920s and 1930s various efforts were made to use the Seashore and other music tests to investigate these popular notions about racial differences. The most easily available comparison was between Negroes and white Americans. Since Negro children were usually behind white children of similar age in school grades, the problem arose 'Should the two populations be compared by age or by grade?' To obviate this difficulty, Dorothy van Alstyne and Emily Osborne (1937) tested children aged from two and a half years to six and a half. Their 264 Negro and 307 white subjects were selected to be as nearly comparable as possible. The children were asked to clap blocks together either in time to patterns produced by Williams' rhythm meter or to reproduce them after listening to each. The Negro children were markedly superior at clapping in time to the apparatus. They were better at simple patterns, and at a slower tempo. Their superiority was less marked at reproducing the patterns and decreased somewhat with age. When schoolchildren and college students are tested, Negro subjects tend to make their best showing on the Seashore rhythm test, and to excel the norms for white populations. However, the differences are usually small (see Kwalwasser, 1955, for a list of various studies and his interpretation of the findings). With the other Seashore tests, white students sometimes make slightly superior scores, but whether this is due to genetic differences, greater test sophistication or to cultural differences, it is hard to say.

Further afield, with the financial support of a benefactor interested in the problem of race crossings in a country with a mixed population, Davenport and Steggerda (1929) undertook a comparative study of full Negro, mulatto and white populations in Jamaica. The Seashore tests were included among the tests they applied. As they decided to discard all results that did not reach the chance level, the numbers remaining in their groups were not very high. On the whole, the Negroes tended to make slightly higher scores than the whites, with the mulattoes coming in between. But the results of the subtests were not very consistent among the age groups and the white people tested were said to be rather inferior and an unrepresentative sample.

It is even more difficult to interpret results obtained by testing non-European peoples who live in very different cultural backgrounds. If lower scores are made by Eskimoes or African villagers, it is difficult to estimate how much should be allowed for 'cognitive' difficulties, i.e. unfamiliarity with English or with the testing situation. Walter Eells (1933) travelled 8,000 miles through the Alaskan winter to remote schools and settlements to test Alaskan Indians, Aleuts and Eskimoes. Their Seashore scores were definitely inferior to the US norms on pitch and memory, though the Eskimoes and Indians were not inferior on intensity. Eell's study was one of the most extensive investigations of primitive cultures with objective tests at the time. Owing to difficulties of travel, he was only partly successful in securing representative samples of the three peoples. Commenting on Eell's study, Cecil Mann of Denver University (1940), pointed out that the difficulty of administering the Seashore tests to primitive people could be appreciated only by one who has tried. He himself had attempted to apply them to over 800 children of Fiji and Indian races. The children experienced such obvious difficulties in understanding and carrying out the directions, that the results had to be discarded as worthless.

Richard Oliver (1932) made a determined effort to test 90 Kikuyu schoolboys with the Seashore test. Their ages ranged from 12 to 24, averaging 19¾. All were learning to sing at school and ten played the harmonium. Before attempting the tests they practised listening to examples and calling out the answers. The average scores they made for Time, Intensity and Rhythm, were well above the American 7th Grade norms. But their average percentile rank for the pitch test was only 40 and for the tonal memory test only 22. Oliver wondered why the scores on these tests were so low, since they had considerable experience of antiphonal singing in their African music where the chorus repeated the part just sung by the soloist.

The poor results may have been due to difficulty in understanding the test instructions. In spite of all his efforts the Kikuyus found the instructions for the tonal memory and consonance tests difficult. Even with a relatively sophisticated student group in India, Parthasarathy (1957) found that Seashore's scores did not follow the pattern set in the American norms, and concluded that local norms would have to be developed before the test could be used in India.

Drake has made particularly careful efforts to test primitive and rural subjects whose exposure to Western music would be minimal. This must have been particularly difficult in the case of his rhythm test, since any music with a steady beat would be relevant. But certainly, the opportunity for training in music and for casually hearing good quality Western music was much less than with most white people. Drake (private communication) collected data on Seminole Indians in Florida, Indians on Guam, natives of the Dutch Antilles, rural Mexicans and a small group of Japanese. He claims that the results he obtained were remarkable in that no real differences appeared. Such variability as was found seemed to be mostly due to error of measurement rather than to any deficiency in any of the ethnic groups. On the other hand, Farnsworth (1931) tested Chinese and Japanese groups who had been living in contact with Western music for varying periods, with the Seashore Pitch and Consonance, and the Kwalwasser Melody and Harmony Tests. Their scores, especially on pitch and harmony, decreased from American standards in inverse relation to the length of time they had been in contact with Western music. White students scored higher on all but the melody test.

Wing (1936, 1948) found little difference between the average scores made by English subjects and those of Jewish, Welsh and German groups. One exception was that 41 German children, aged 12–13, seemed better able to do his appreciation tests than English children. He believed this might be due to the greater amount of first-rate music which the Berlin child could hear, and perhaps to the fact that since so many of the great composers were German, it made some appeal to their national pride.

Both Sanderson (1933) and Witherson (1935) found that Jewish children did on average make rather higher scores on the K–D Test than did some other national and racial groups. Sward (1933), employing the more valid Drake Memory Test, as well as some of the Seashore and K–D Tests, came to the conclusion that 200 Jewish children were only slightly superior to non-Jewish subjects. He contrasted this very small difference in talent with the very much greater achievement of Jews as professional musicians. About 50% of violinists in American symphony orchestras were at that time of Jewish origin. 10% of American composers were of Jewish extraction. If the proportion of the highly talented is in fact no higher among Jews, there must be much talent among Gentiles which does not find expression in music. The professional eminence attained by the Jews may sometimes have only been achieved at the cost of a more intense effort than many Gentiles would be willing to expend.

While the current tendency seems to be to interpret such racial differences as are found in terms of differences in educational and cultural background, we cannot be sure that there are not also genetic variations. Negroes may be better at the rhythmic aspects of music because less sophisticated instruments were developed in Africa than in Europe, or because as children they are not discouraged from clapping their hands to music and beating on drums. But there might also be some inherited difference, just as there is in the case of blood groups among the populations of different parts of the world.

With the spread of Western music to all parts of the world, it might be useful to develop local norms for British and American musical ability tests. Comparable types of tests could also be produced, based on Oriental or African music.

XV

The Effects of the Home and Social Environment

The influence on musical ability of a child's home environment (which will be taken to include any substitute home in which the child grows up) is almost literally 'incalculable'. Some estimate can, however, be made by comparing children from homes where both parents are performers and where good music is listened to frequently with those who come from homes where neither parent plays and where the only music to be heard is 'pop'.

There seems to be some association between musical environment and the Seashore results, though there are just as many with superior talent from 'C' as from 'A' homes. It is quite possible that any increase in stimulation beyond a certain point may make no further difference to the child's ability. Stanton herself interpreted the higher proportion of individuals with superior profiles who had enjoyed good musical backgrounds as children in terms of their having inherited their talent from parents who, because they were musical themselves, provided a considerable amount of music in the home. Because of their remoteness from music, the Seashore tests are probably not particularly susceptible to the direct effects of general exposure to music in the environment.

Wing (1948) who wished to establish that ability to perform his tests 'was not unduly influenced by opportunity to hear music' collected some interesting data on the effects on a child's test scores of the music played by other people in his home. He found that there was a significant difference between the scores made by 333 boys who had music at home when their parents played but *not* when persons other than their parents played. He also showed that there was only a very moderate association between interest in music and ability to perform his tests (an approximate correlation was about ·30). In any case, as far as awakening the child's interest was concerned, parental playing was very little different from the playing of others. He concluded, therefore, that the most likely explanation of the association between the child's ability and parental playing was that the child's ability had been inherited.

This may be the true reason, but it is not the only possible one. Parental playing may be a more potent influence on the child merely because it has gone on for much longer or because the parents have more prestige in the eyes of the child. Wing rejects the latter interpretation, though it could be argued that adolescent boys might have been more influenced by the playing of their parents than by the practising of their sisters. Wing does not state how the group of 'non-parents' was made up, but it may have included a number of lodgers (Wing – private communication) as well as other blood relatives, siblings with perhaps some grandparents or uncles and aunts. (In so far as the observed significant difference was between the parents' playing and the playing of less close relatives, as opposed to unrelated strangers, that would tend to strengthen the hereditarian argument.) Wing's percentages for adolescents who give up playing and for adults who have received music lessons (see p. 34 above) do confirm that parental playing is a rough indication of musical ability. However, almost 25% even of Training College students with above average ability had never received lessons nor tried to play on their own. Moreover, many of the parents who played may have had little real ability.

Wing excluded separate consideration of the influence of radio music. Broadcast music is now, however, likely to be a much more potent influence than in the late 1930s when Wing was collecting his data. There is little real evidence of the effects of broadcast music on musical ability. The more valid tests were perhaps standardised a little too late to show whether there has been any rise in musical ability analogous to the increase in IQ which tends to occur when children are removed to a more highly stimulating intellectual environment. It is reasonable to suppose that broadcasting has tended to make the musical environment more uniform since nowadays even the child from a very poor home can see and hear orchestral instruments at a very early age if he so wishes. No doubt the greatest effect will be on children with a high degree of talent and interest since they will learn more readily through their ears. As stated in *Teaching Music in the Schools* (Min. of Education, 1960)

> 'When a child comes to school he normally brings with him a considerable variety of musical experience. Much of this will doubtless have come from sound radio and television programmes ranging in suitability from such series as *Listen with Mother* to material of a more sophisticated character preferred by the older members of the family.'

However, much broadcast music doubtlessly falls on inattentive ears.

Even in 1934 Constant Lambert complained 'Never has there been so much music-making and so little musical experience of a vital order.' The child trying to complete his homework against a background of music may seem to be receiving training in *not* attending to music. A broadcast concert may often become something narrowly sandwiched between finishing homework and getting ready for bed. The danger of broadcast music is that much-played pieces may become overfamiliar, or heard superficially so often as to dull the appetite for any real appreciation. 'The objection to a constant broadcast stream of light music is not that the music is light' (if music is to serve merely as a background, then 'serious' music is too good for that purpose), 'but that it is hypnotic and its associations maintain a constant pattern of mild titillation' (Grey Walter, 1953, p. 47).

The child whose parents set an example of good listening is likely to value music. However, if the child's interest in music can be awakened at school or by private music lessons, the radio and the record player provide great opportunities for its development. So long as the rest of the family is willing to have even such programmes as *Your Hundred Best Tunes* switched on and are not making a noise or talking, the child can still listen, even if his father is filling in his football pools coupon, his mother knitting, and his sister reading.

William Kirkpatrick (1965) of the University of South California, found a strong relationship between the singing ability of over 100 five-year-old children and their home environment. From recordings of their repertoire of songs, he classified as 'singers' children who could sing 90% of the notes correctly without change of key, as 'partial singers' those who could sing 75–89% of the notes with only a few changes of key.

He regarded as 'non-singers' the children who could sing less than 75% of the notes correctly or who sang without any established tonality. Few non-singers came from homes classified as excellent or good from the point of view of the musical environment. No singers came from musically poor homes. A few children, possibly because of genetic endowment, seemed unaffected by their musical environment. The influence of television and record-players on singing ability was negligible. A significant relationship existed between singing ability and the following environmental factors: mothers who sang to and with their children, aid from other adults other than parents, family singing and playing, and parents with a musical background. Older brothers and sisters and attendance at a nursery school had a lesser influence on children's singing ability.

An interesting study of parental attitudes to their children having music lessons was carried out by Winifred Graves (1947). An American clinical psychologist who had herself at one time taught the piano, she lived in a community where the giving of music lessons was a flourishing industry. She tried to find the answer to questions such as whether the child who was having lessons might gain satisfaction in playing because of native ability and a real pleasure in music, or whether he tried to excel in music to offset feelings of insecurity in other areas; and whether parental attitudes towards studying music might reflect genuine interest in the child, or a desire for social prestige due to the child's success, or unfulfilled parental ambitions projected upon the child. She compared two groups of children aged between nine and seventeen, matched in age, intelligence, sex and school grades and differing only in that one group was taking music lessons. She found that the children taking lessons were *not* victims of parental over-direction. On the contrary, having music lessons was associated with being in harmony with one's family, not having lessons with being in conflict with one's parents' ideals and actions, one's own ideals and one's friends' behaviour, as measured by the Spencer Conflict Score. There was no significant difference between the two groups of children in their attitude to music, except when the child's belief concerning whether his parents wanted him to have lessons was taken into account. The parents whose children were learning to play estimated children's joy in playing as much greater than did the control group's parents. Of the 25 children of the experimental group where data were collected from both parents, only 3 fathers had not had music lessons themselves, whereas 17 of the control group's fathers had not learned music. The control group's parents had not liked playing, but the experimental group's parents had enjoyed playing. Graves later carried out a more extensive study, using 71 pairs of adolescent children, but a full report has not been published. The findings, as summarised in *Dissertation Abstracts*, indicated that children whose parents provide them with lessons are rather better adjusted emotionally than those who do not receive lessons.

As part of her investigation of parents and children (see p. 121), Shuter tried to determine the effect of various environmental factors on the children's musical ability.

She compared the children's MQs with a listening score. This was derived from questionnaire data on the amount of concert-going and listening to music on the record-player and to broadcast music. She considered it best to accept her subjects' answers as given, though it was difficult to judge how literally they had taken the instruction to include listening to the radio only when *not* doing something else. When computing the scores, a weighting of 3 for classical, 2 for ballet and opera, and 1 for other types of music was applied. The listening scores correlated only ·177 with MQ. Inspection of her results showed that the lower listening scores were generally associated with lower MQ levels. This seemed particularly to be expected with grammar school children, as the pressure of school work and the wide variety of other interests available to them, could easily lead to the less talented neglecting music. On the other hand, there was less agreement at the higher levels of music. This may have been due to talent not necessarily being accompanied by interest in music or because many musical individuals may prefer to devote the time they have available for music to performing rather than to listening. (Shuter's results comparing the children's musical activity with MQ are discussed on p. 170.)

Shuter also made a detailed analysis of the parents' questionnaire answers, tabulated against their children's musical level. An index of the child's musical level was produced by adding the marks for musical knowledge, listening and activity to their Wing scores. On the basis of this total index of musicalness, they were divided into three groups.

The parents of the bottom group were, on the whole, considerably less active musically than those of the other two groups. Within the listening groups of factors home listening seemed more important than concert-going, perhaps because the subjects lived a fair distance from London, at least from the point of view of parents many of whom may have had children younger than those who took part in the experiment. But only in the case of listening to classical records did the association between parental listening and child's musical level reach a statistically significant level. On the other hand, present parental playing and music lessons were related to the children's musical level to a highly significant degree. The hereditarian would argue that the parents who had had music lessons and kept up their playing were those who had talent and that their children had inherited their gifts. But equally the environmentalist could hold that the parents' playing had contributed to raising the children's musical level and that parents who have themselves had music lessons are more likely to encourage their children to learn.

Similar arguments could be applied to Holmstrom's findings (Holmstrom, 1963). He compared the scores made on his version of Wing tests 1 to 3 by children from musically good homes with those from musically poor environments. The differences in their scores remained statistically significant in favour of the children from the musical homes even when the effects of interest in music and intelligence had been removed by statistical procedures.

In Holmstrom's study the 'good' home was characterised by having a radio and one or more 'noble' musical instruments. At least one member of the family played every day and listened frequently to music on the wireless. The musically 'poor' home had no musical instruments and no member of the family played and seldom or never sang.

Edward Rainbow's (1965) researches threw a little more light on this question. Using a more detailed questionnaire than Shuter, he obtained a separate index of 'home enrichment' and 'participation in music by relatives'. His subjects were pupils at the Laboratory School of the State University of Iowa. Their IQs averaged about 116 – roughly similar to the intellectual level of Shuter's grammar school children. 91 of Rainbow's subjects were at the elementary school stage (average age $10\frac{3}{4}$), 112 at junior high school (aged about $13\frac{1}{2}$) and 88 at high school with an average age of $16\frac{1}{4}$. In arriving at a score for home enrichment, Rainbow gave one-half point if there was a television set in the home, one point for a radio or record-player, two points if the parents had encouraged the child to belong to a music group, or if his brothers or sisters played a musical instrument or if the family, as a group, listened to music. Three points were awarded if the family performed music together. In the case of participation in music by relatives, Rainbow awarded two points if either parent played (or had played) a musical instrument or sang, and six points if both parents played or sang. One point was given for each grandparent and a half-point for any uncle or aunt who was reported to be musically active.

Evidently many of the parents who did not play themselves encouraged their children to take part in music – judging by the quite low correlations between home enrichment and participation in music by relatives (highest $r = \cdot45$). The effect of home enrichment on the scores of Seashore's pitch, memory and rhythm tests and of Drake's memory test was quite small, the test most affected being the Seashore memory test (highest $r = \cdot34$). The correlation with the teachers' estimates of their pupils' potential capacity for music was also low. Within the three age groups, home enrichment did differentiate between the 20% of the children judged by their teachers to be the most musical and the 60% rated average in musical potential, and also between the lowest 20% and the average group. Home enrichment seemed to have a positive effect ($r = \cdot415$) on the youngest children's interest in music and, as we might expect, less influence on the interest taken by the oldest group in music.

In Rainbow's results, performance of the music tests, and music potential as estimated by the teachers, appeared to be wholly independent of participation in music by relatives among both the youngest and eldest children. The lowest 20% among the junior high school children received a very low average mark for relatives' participation. This led to only a small positive correlation with the musical ability tests. The difference between this finding and Shuter's may be partly accounted for by Rainbow taking into account *any* playing or singing by the relatives. However, Wing, too, did not assess the amount of parental playing and found, like Shuter, an appreciable association between their musical activity and their Wing scores.

THE SOCIO-ECONOMIC STATUS OF THE HOME

How far, we may ask, does the socio-economic status of the home affect the children's musical ability? Is any superior performance of a music test to be attributed to the generally more stimulating environment of a 'better class' home? In considering this question we must bear in mind the positive relationship between socio-economic status and intelligence. The correlation of father's occupational level with child's IQ has been consistently found to be about ·35 (Vernon, 1960). If the higher class child does better at a musical test, it may be partly because he is more intelligent. Intelligence is likely to have been a factor in the early studies of Burt (1909) and Valentine (1962) where children from preparatory schools were compared with elementary school children. However, English children attending preparatory schools in the first two decades of this century were very likely to have come from homes of superior socio-economic status.

Burt found that 13 preparatory school boys (aged 12:6 to 13:6) were better at the discrimination of pitch than 30 elementary school boys. Half of the latter were choristers and many were learning some musical instrument, while 5 out of the 13 preparatory school boys neither sang nor played and the whole group had received much less training in music.

At a preparatory girls' school in which nearly every girl over seven learned some musical instrument and all of whom heard good music fairly often, Valentine found that the girls developed the power of discriminating between concords and discords as much as three years earlier than did elementary school children. By the age of nine they gave an order of preference for musical intervals almost identical with that given by adults. Writing in 1962, Valentine himself expressed the view 'it is just possible that superiority in general intelligence may have helped these preparatory school girls of only 6, 7 or 8 in understanding what was required of them'. He wondered that 'whether far different results from mine would be found nowadays with children in our Junior or even Infant schools in view of the much more frequent hearing by children of music owing to the radio and the greater place given to music in the Primary schools than was the case when I did my experiments over 40 years ago'.

The children from better class homes have more opportunities for music lessons. Gilbert (1942) suggested that this may improve their music test scores. When he re-analysed his data (see p. 88), classifying by socio-economic status instead of by sex, he found that the highest college scored 209·8 points on the K–D tests, the middle one 203·7 and the lowest 200·6. However, the higher the socio-economic status of the college, the greater the percentage of students who had received music lessons. When only untrained students were considered, the mean score (197·0) was about the same at each level. Gilbert, therefore, was inclined to conclude that the better results of the trained group were due to their having enjoyed opportunities for music lessons. He rather oversimplified the picture, however, by neglecting to consider the probability that many of the students who had learned music had musical parents from whom they may well have inherited their talent.

In his earlier thesis, Wing (1936) reported that the average scores of elementary and secondary school boys were almost identical except on tests requiring aesthetic choice. He thought the superiority of the secondary school boys might be due to their having heard more good music at home; less time was spent on music at their school compared with the elementary school.

Type of school attended is obviously only a rough indication of socio-economic status. In the following studies the classifications were based on fathers' occupational level.

Parker (1961) tested over 1,000 Kansas high school children with the Gaston test of musicality and with the Wing appreciation tests. With intelligence and the Gaston score held constant, the correlation between the Wing appreciation tests and socio-economic status was zero. Shuter (1964) investigated how far the Wing scores of 189 junior musicians of the Royal Marines School of Music were related to their fathers' occupational level as stated on the entry forms filled in by the boys. When the two top grades of Wing scores were taken together, more than twice as many came from the highest social categories than from the lowest. However, 8 boys out of 18 boys whose fathers were unskilled manual workers had above average talent. The differences were not statistically significant partly perhaps because of the absence of representatives from the highest socio-economic categories. Moreover, no information was available about the early home background or occupation of the boys' mothers. This might be particularly important with a subject like music.

Rainbow (1965) assessed socio-economic status on the basis of the education as well as the occupation of the head of the household. He found that the correlation between socio-economic status and home enrichment was about ·3. This would confirm the everyday observation that there is a tendency for musical activity in the home to be related to socio-economic status but that higher social status and a regard for music do not always go together – the children of a Welsh miner may be in a more musically stimulating environment than the offspring of a Midland industrialist.

The correlations between socio-economic status and the music tests were low. But when Rainbow analysed his results by a multiple regression technique, socio-economic status was shown to be contributing to a statistically significant extent to the children's musical aptitude as estimated by their teachers, both in the case of the elementary school and of the total group. We may wonder, however, whether the teachers were in some way influenced in their assessments by factors connected with the socio-economic status of the pupils.

As we might expect, children from a good socio-economic background seem to know more songs and be able to recognise more tunes than children from less favoured homes. At least this was the conclusion reached by the writer from an investigation carried out in four Glasgow schools (Jamieson, 1951). She also found, however, that adverse home conditions could be partly counteracted by really enthusiastic efforts to foster music on the part of the school. Comparatively high scores were made by the children, especially the girls, who attended a school in a working class district where the headmaster was particularly keen on music.

Again, both Rogers (1956) and Baumann (1960) found a tendency for higher class children to like classical music more frequently than did children from homes in the lower social categories. However, at least with their (American) subjects, in all social classes popular music was the type most enjoyed.

What can we conclude from the research reviewed in this chapter? A musically stimulating home is certainly likely to help children to make the best use of whatever potential talent they may happen to possess. It would be over-optimistic, however, to hope that parents could substantially improve the ability, as opposed to the taste, with which their children have been endowed.

XVI

Effects of Specific Practice and Instruction

Can pitch discrimination, performance on musical ability tests or such activities as clapping in time to music or singing intervals and phrases be improved by coaching and practice?

PITCH DISCRIMINATION

Most of the research relating to the possibilities of improving sensory capacities has been devoted to pitch discrimination. Seashore believed that each individual has a maximum potential power for sensory discrimination, which, being determined by the inherited efficiency of the ear, was reached early in childhood and could not be improved by environmental influences. In his book (1938) he wrote:

> It seems probable that just as the physical eye of the child at the age of three is as keen as it will ever be, so the pitch sensitiveness in the ear reaches its maximum very early. . . . The physiological limit for hearing pitch does not improve with training. Training, like maturation, results in the conscious recognition of the nature of the pitch, its meaning, and the development of habits of use in musical operations.

Seashore repeatedly cautioned the user of his tests that, although the aim was to determine the individual's maximum capacity, the result might actually fall short of this ideal, indicating a cognitive limit. Failure to reach the physiological limit was ascribed to such factors as lack of concentration or motivation, or a failure to understand the instructions. Any improvement in scores on later testing was explained by the elimination of such factors or perhaps to the choice of a better work method.

Seashore's position was challenged by Ruth Wyatt of Northwestern University, Illinois. In her monograph (1945) she presents a detailed and critical review of the literature and describes an experiment of her own, where considerable improvements in pitch discrimination were achieved by her subjects.

A wide variety of procedures has been used in the study of the improvability of pitch discrimination, ranging from merely repeating the tests with little or no change of method to intensive remedial training. Techniques of remedial training have included the following:

1. Telling the subjects whether or not their replies are correct;
2. demonstrating with foreknowledge the correct answer;
3. drill in discrimination or recognition of intervals on the piano;
4. vocal matching of tones, phrases or scales;
5. other techniques of illustration or explanation.

EFFECTS OF REPETITION

Seashore (1919) quoted an experiment by H. S. Buffum which he interpreted as showing that pitch discrimination is unaffected by training. Twenty-eight fourteen-year-old pupils were given 40 minutes of 'specific and intensive practice' for 20 successive days. Only two were found to have improved. However, the children were not given any help or real remedial training, during what must have seemed to them a very boring task.

Frances Wright (1928) tested 24 adult music college students with the Seashore pitch test every day for a week. The mean scores for the group improved only slightly. However, Wyatt pointed out that, while the six highest subjects had made slightly lower scores on the last day, the six lowest students had improved by small but steady daily increases (except on the fourth day), from a score of 81 to 87. There was of course more headroom for improvement among the lower scores; the more talented students may have become bored.

One of Teplov's subjects improved considerably during the course of four testings without any special training. As she had never learned music, the task of comparing two tones was new to her and what Seashore would have called cognitively difficult. But two subjects with long musical experience also improved just by hearing others being tested.

It would thus seem that pitch discrimination can to some extent be improved by practice alone.

EFFECTS OF REMEDIAL TRAINING

As long ago as 1903, Whipple reported on the pitch discrimination of a woman who had never been able to sing in tune and who could not detect a change of a semitone in a familiar melody. By 'systematic drill and coaching', Whipple succeeded in reducing her pitch discrimination threshold from 12 cps at a standard of 250 cps to 2·8 cps. This improvement did not, however, transfer very effectively to piano tones.

Smith (1914) gave the lowest 106 of 476 child subjects 'special assistance' to help them to 'distinguish different tone qualities and to form the right habits of attention'. For the 71 boys, the average threshold was reduced from 17·3 to 9·8 cps while that for the 35 girls dropped from 17·7 to 7·8 cps. Smith also gave two pitch discrimination tests to 200 adults. The poorest 54 subjects then received 'personal instruction'. They were given explanation based on the diagnosis of individual difficulties. Forty-seven improved rapidly. After instruction 28 had thresholds between 3 and 0·5 cps. Those with high thresholds (12 cps or more) had been reduced from 51% to 6%.

Following Seashore, Smith interpreted the improvements as being due to the removal of 'cognitive' factors. Wyatt, however, thought that improvements in work methods had also taken place, since analysis of the subject's introspections showed that factors such as producing the tones vocally or subvocally and learning the optimal adjustment of attention were of importance.

More definite remedial methods were used by Robert Seashore (1935). He selected 12 adults whose PR on the Seashore pitch test was 12 or less. A beat frequency oscillator was used during the training. His procedure included: (1) giving demonstrations of easily noticeable differences in pitch, having told the subjects beforehand what was to come each time, and (2) telling the subjects whether each judgment was right or wrong. After the first period, most of the time was spent practising slightly below the most recently determined threshold – thought to be the most efficient level for the student to concentrate on at any given time. The total time devoted to the training ranged from 3 to 8 hours, averaging 5–6 hours, given in weekly periods of 45 minutes. This relatively short amount of training resulted in 10 subjects improving, 7 achieving thresholds of 3 cps or less. Three of them improved their PRs on the Seashore to 91, 81 and 70.

Worthwhile improvements were reported by Connette (1941) after giving 23 adults only five days of individual practice. Like R. Seashore's subjects, they were told their results after each response and given demonstrations with foreknowledge of the correct answers. The subjects in the upper half on initial scores improved by 29%.

Teplov and Borissova (1957) also succeeded in improving the pitch discrimination of some of their subjects by giving knowledge of results. If the subject made a mistake he was told the correct answer and asked to listen to the two tones again. The improvement occurred very rapidly, sometimes in one training session. Teplov concluded that in only a few hours of work, thresholds could be reduced by half or more.

Capurso (1934) experimented to see whether training in interval recognition, in addition to practice in discriminating the higher of two tones sounded on tuning forks, would improve Seashore pitch scores. After testing 58 adult music students, he chose four with the highest and three with the lowest scores as an experimental group. The control group consisted of three from the highest and three from the lowest scorers. Capurso's method of instruction in recognition of intervals was to try to get his subjects to associate piano intervals with a 'mood word', such as 'tumult', 'comfort' and with other auditory stimuli. For instance, the fifth was to be associated with the ringing of chimes. This part of the training was very different from the discrimination of the small intervals used in the Seashore test. After half an hour's practice on alternate days for seven weeks, only the lowest three in the experimental group had improved their scores significantly. One increase was spectacular – from a score of 50 (equal to PR 3) to 90 (PR 96). Capurso also carried out a similar experiment over six months with two very pitch deficient subjects. One improved only from a score of 53 (PR 3) to 71 (PR 17). The other student improved her PR from 6 to 94 and learned to sing scales without the piano. Most of the improvement occurred by the time she had formed associations for all the intervals and before training with the tuning fork was begun. As can be seen from a verbal report which Capurso reproduces, her violin playing also improved.

> When I practised (violin) I would become discouraged and would give up in utter despair. If, for instance, I played a tone on the violin and tried to compare it on the piano, I could not tell whether they were the same or not. . . . When I came to the University I was ashamed to say I had been taking lessons for so long a time and accomplished so little. . . . Now, having had some training in pitch, I am beginning to hear the tones and am being [*sic*] able to discriminate especially well in the low register which was originally the harder for me to hear.

Of special relevance to the problem of tone-deaf children is the study carried out by Wolner and Pyle (1933). They sought out the most pitch deficient children in fifth to seventh grades in all the schools of Detroit.

The seven children they selected for training were extreme cases: they could not sing and could not discriminate pitch differences as large as 30 cps at a standard of 423. Daily individual training sessions of 20 minutes were given for 81 days. At the beginning the child practised singing back one note played on the piano. When one note had been mastered, intervals and short phrases were given, later progressing to scales. Drill at discriminating intervals was given; whenever a misjudgment occurred the child had to go back and sing the interval. The child also sang the larger differences (30, 23 and 17 cycles) with the tuning forks.

The children were repeatedly told to listen to the notes first, to try to get a mental picture of them, and then to reproduce them as accurately as possible. All seven learned to discriminate perfectly the intervals of octaves, fifths, thirds, whole tones and semitones over a range of four octaves. Each pupil improved noticeably in ability to sing. At the end of the experiment, one could sing several songs without any trace of pitch deficiency, and major and minor scales. Another could sing scales, intervals and the music of a song without the words. Two could sing scales and intervals. The other three had not learned to sing perfectly, but had improved tremendously. Considerable patience and effort were needed to keep the children highly motivated. The patterns of learning which occurred were similar to those found in experiments on the learning of skills. The greatest difficulty was found with differences of 30 cps and 23 cps. Once these were mastered, a spurt of improvement followed for several more difficult increments, only to be followed by another period of seemingly arrested progress at an increment which required another delay before it could be mastered and progress continued. It took one boy four weeks to conquer the 30 cps fork. After this he passed the 23, 17, 12 and 8 forks with comparative ease. The 5 cycle fork took a week to master. The 3 cps fork took two weeks. From this he went on to the 2 fork and remained at it for a whole week. Here he had a little trouble. At both the 2 and 3 cycle levels the influence of intensity was apparent to a marked degree. Upon completing the 2 cycle fork successfully, he had no difficulty with the 1 and 0·5 cycle forks.

For her own research Wyatt chose as subjects eight adult music students who were fairly proficient in singing, and eight students in Liberal Arts, most of whom had experienced difficulties with pitch. Even the most able could not sing a melody if the person next to him was singing a different part. To establish their initial level of performance, they were tested at least twice, with the B form of Seashore's pitch test (see p. 280) and a test of Wyatt's own. Both measured down to pitch differences of 0·5 cps at 500 cps.

The subjects were given twelve periods of 50 minutes individual training. To train pitch intonation each student was asked to listen to a tone on an oscillator and to try to sing it into the microphone immediately after the tone had ceased. If he was unsuccessful, he was asked to try to match his voice to a piano note or to a note sung by the experimenter herself. Once the single tone could be reproduced correctly, the student learned whole and half-tones above and below the standard. The semitones were associated with d^1t and td^1. A chromatic stroboscope was used so that the subjects could see as well as hear how well the notes they sang matched the standard. The pattern produced on the stroboscope remained stationary if the tone was exactly in tune but moved to the right if it was sharp and left if it was flat.

The training in pitch discrimination was begun at an increment slightly smaller than one at which the subject had achieved 40 correct answers. If the subject gave a wrong answer, he was told about it immediately. The interval was demonstrated to him. If he could not discriminate whole and half tones, he sang with the tones. If he claimed both tones sounded alike, he was asked to verify the difference by looking at the stroboscope. He was told to try to form a clear auditory image of the standard. Forty perfect answers had to be given at each increment before going on to the next most difficult. One subject, a professional piano teacher, commented that 'a difference of 4 or 5 cycles which formerly gave me uncertain moments and sounded like one tone, now seems very far apart, and it is no effort to distinguish the direction of the pitch difference'.

The postural attitudes adopted by the subjects seemed to be significant, but there were individual differences and none was optimal for all the subjects. Two subjects had a larger percentage of correct judgments when they seemed most relaxed, when they were looking around the room, or toying with their hair. Others did better when they concentrated. The most helpful procedures were those that encouraged the use of auditory imagery and motor participation.

At the end of training, statistically significant improvements had been made both by the music and non-music groups. The Seashore score of the music students had increased by an average of 7·75 points and their Wyatt score by 12·25. Of the six who were initially 'poor' according to Seashore's grades, two had become 'good', one 'excellent' and two 'superior'. The non-music group improved their Seashore scores by 4·50, on average, and their Wyatt scores by 14·95 points. The music students greatly improved their pitch discrimination of oscillator tones at each level of difficulty, the non-music students at every level down to 8 and 5 cps.

Wyatt certainly showed that pitch discrimination can be improved by really efficient and enthusiastic remedial training adapted to the needs of the individual. However, there do seem to be some cases of tone deafness which can be improved to only a very limited extent. Though Stumpf (1883) succeeded in improving the judgment of piano intervals in most of his subjects, he failed with a woman who seemed to hold a record for tone deafness. She could hear a difference between notes two octaves apart but could not say which was the higher. Teplov was able to reduce one woman's threshold from 226 cents (over one tone) to 94 cents (less than a semitone) but further progress proved impossible. When the second note was higher, her discrimination was better than when it was lower. Teplov concluded that such people had difficulty in distinguishing between the pitch and the timbre of sounds.

Tomatis seems to have found a means of improving tone deafness by controlling the harmonics of sounds during training. While he was studying on behalf of the French Air Force the effects on hearing of working close to jet engines, he observed that people who had lost the ability to hear high notes did not use them in their speech. Not only did he find that the way in which people speak is affected by the way in which they hear themselves, but he discovered that one ear must be 'dominant' to provide correct feedback to the brain. In right-handed people, the right ear is the directive one, in the left-handed, the left. Trouble arises in people who have no clearly dominant ear. In 1961, Tomatis demonstrated in London and Edinburgh how people can be trained to sing in tune by correcting electronically the faulty response of their directive ear. Among his patients was a Benedictine monk who 'sang terribly'. After a course of listening through earphones to his own voice distorted electronically to compensate for the shortcomings of his ear, the monk learned to sing in tune.

The use of modern electronic equipment to help pitch deficient people has also been developed in Russia and the United States. A. N. Leontiev (1957) trained tone-deaf children to attune their voices to the pitch of sounds fed as continuous notes into earphones. As soon as the child began to sing, the sound was switched off and he continued to sing independently. Later, intervals of up to six seconds were introduced between the sounds and the child's response. All the children improved. Leontiev also attempted to train them to reproduce very simple melodies, but did not give details of the results.

Because of the value of individual coaching in the training of pitch discrimination, the American psychologist, B. F. Skinner, has used a 'teaching machine' for this purpose (see Chapter XXVI). Beginning with coarse discriminations, his tone-deaf child subjects were gradually brought to the point where they could discriminate tones and semitones.

The implications of the studies reviewed above for the education of the tone deaf will be further discussed in Chapter XXV.

Singing Ability. Granted that gross pitch deficiency can be improved by remedial training, the question then arises 'What has been achieved that is musically useful?' The findings of Kalmus (see p. 134) and of Fieldhouse (see p. 197) suggest that tonal memory is an even more important factor in tone deafness than pitch discrimination. In a study at Durham, Pollock (1950) found 45 'monotones' were significantly inferior to 45 normal children in both Wing's memory test and the whole Wing battery. Unfortunately, no information seems to be available about the subsequent history of any of the pitch deficient subjects in the studies discussed above. It would be of great interest to know whether any all-round improvement in their musical ability eventually took place. It seems reasonable to suppose that those who learned to sing one or two songs with perfect intonation would be able to increase their repertoire. How far they could progress to other activities (such as holding a part against an upper voice or learning a pitch-variable instrument) is less certain.

But at least the subjects who learned to sing a scale and musical intervals in tune, without the piano, would seem to have acquired the absolute essential basic skill required for further progress with aural training. They would now be able to undertake a course of ear-training such as described by Annie Lawton (1933) with some hope of being able to master the first few exercises in pitch. Lawton, a teacher of considerable experience at what is now the Royal Scottish Academy of Music, intended her course primarily for adult piano students who were weak at aural work. But it could be used with a teacher. Her first pitch exercise requires the student to sing up and down the scale, sounding each note on the piano before singing it. Even this would prove difficult to many pitch deficient subjects before remedial training. In the second exercise the student has to sound the key-note on the piano, then sing up the scale, testing the accuracy by playing each note after it is sung. This cannot be done without a clear memory for the relative pitch of the second, third, fourth, etc., notes. But if this exercise and succeeding ones on the tonic chord can be mastered – and Wolner and Pyle's subjects should have been able to perform them successfully after their course of training – then the way would seem to be open to sight reading three or four notes (given the tonic) in major keys, then in minor keys, two-part melodies and for exercises in recognition of cadences and in simple dictation. Lawton, like Wyatt, stressed the importance of careful listening to establish a clear mental concept that will remain in the memory and can later be reproduced vocally and eventually by 'silent mental repetition' (probably accompanied by sub-vocal movements).

The importance of being able to sing in tune is here emphasised because, from the point of view of improving pitch defects, it makes practice away from the piano very much easier (and provides overt evidence of the effects of training). How far such exercises would lead to an overall improvement in musical ability has never to the writer's knowledge been systematically studied. There is, however, evidence from carefully carried out experiments in the USA that the singing of single notes, intervals and phrases by young children can be improved by training and that the effects of the training persist.

Thus, Arthur Jersild and Sylvia Bienstock (1931) trained 18 children (average age 3:2) in the singing of single notes and of intervals during 40 ten-minute sessions spread over six months. A significant difference was found between the final scores of the experimental subjects and those of 18 other children paired with them. When the same children were retested two years later (Jersild and Bienstock, 1934) the trained group had retained a reliable superiority over the control group.

The initial scores made by the three-, four- and five-year-old children who took part in Updegraff, Heileger and Learned's investigation (1938) have been quoted in Chapter VI. After 15 days, after 30 days, and in the case of the five-year-olds, after 40 days, the children were retested. Definite, consistent improvement was found among the children given training. The control group improved slightly, if at all. In the case of the three-year-olds, the curve of improvement was slightly steeper between the fifteenth and thirtieth periods than earlier. The same was true of the four-year-olds in the phrase test, but not in the easier single tones and interval tests. The five-year-olds improved more rapidly during the first 15 days. The improvement at all ages continued throughout the period. One of the five-year-old boys in the experimental group was at the beginning 'practically a monotone'. He improved during the course of training, though his record materially affected the mean scores of the group.

The interest of the experimental group in the actual training procedure remained high throughout Jersild and Bienstock's experiment, though no systematic method of estimating interest was used. During the Updegraff, Heiliger and Learned investigation, both the experimental and control children were carefully observed during periods of musical activities at the school. The degree of interest and voluntary participation shown by the children was carefully recorded on an observation blank and independent ratings made by the experimenters and teachers. At every age the children of the experimental groups showed increasing interest in the musical activities of the school. The interest scores of the control group reflected little or no change in attitude. By the end of the training period the difference between the two groups was statistically significant. The normal music training at the nursery school attended by these children, though not inferior to average, may have been below their capacity to respond, thus allowing considerable scope for the trained group to show improvements. It is also possible that the gains made were in the nature of accelerated development. If they lasted, or were boosted by further training, the final result might be that the children would reach musical maturity at an earlier age than normal, with their ultimate maximum of development remaining unchanged.

Improvements in singing can also be produced without recourse to individual training. Leon Culpepper (1961) at the George Peabody College for Teachers gave half an hour daily practice in singing with records for four and a half months. Most of the children improved. Certain stages of development in learning to sing in tune were discernible and Culpepper claimed that the level of development could be determined by finding interval between the tonic and the note substituted for the tonic. Children whose singing was 96–100% defective sang a note a sixth below the tonic, those whose singing was 80–90% defective sang a note one-fourth below and those whose singing was 60–70% defective sang a second below the tonic. Those children who sang less than 50% of the notes out of tune could sing the tonic correctly most of the time.

Robert Smith (1963) gave group training to 20 three-year-old and 21 four-year-old children at the University of Illinois Child Development Laboratory. Each day for two 16-week terms they received 15- to 20-minute periods of practice in singing. All improved, especially on the lower register, C to A.

Although such training does seem to lead to improvements in singing, individual differences remain. Williams (Williams *et al.*, 1933) found a considerable range of ability to sing songs which had been learnt during a whole session among children who had received a similar amount of intensive training.

EFFECTS ON MUSICAL ABILITY TESTS

If the experimental groups mentioned above could have been given further training, for example in recognising which note in a short phrase had been changed on a second playing, it does seem likely that such long-term practice would tend to improve performance of tests such as Drake's and Wing's Memory Test.

Some empirical evidence is available on the effects of rather short-term specific coaching and of aural training on tests of musical ability based on musical material.

Drake (1945) gave both forms of his Musical Memory test to 14 college students (presumably specialising in music) and to 58 psychology students. After the experimental group had attended an ear-training class for a term, both groups were retested. The trained group had improved their mean score from 27·72 (number of mistakes) giving a PR of about 50, to 19·43, which is equivalent to a PR of 80 from the norms for 'Music Students'. The 58 control students had reduced their average number of mistakes from 46·24 to 42·02, representing an increase of PR of 10 points, from 50 to 60, on the norms for 'non-music students'. The increased improvement due to training, as opposed to mere retesting, was not significant statistically. It is a pity that the two groups were not more comparable in past experience. As the control group had evidently much less previous training (though equal in initial percentile rank on the different set of norms) it is possible that if the experimental subjects had been chosen from them, the scope for improvement by ear-training might have been much greater.

Edwin Gordon (1961) investigated the effects of specific coaching on the performance by 14- to 15-year-old children of the two Drake tests. He chose 10 children whose initial PR ranged from between 50 and 75 and 10 whose PR ranged from 1 to 36, on the Memory test. Five from each 10 were selected to serve as the experimental group. Seventeen half-hourly instruction periods were devoted to the practice of musical phrases similar to those used in Drake's Memory Test. The subjects were taught how to listen for the different types of changes in the phrases. Neither the control nor the experimental group was receiving outside musical instruction. Both groups were retested at the end of the experiment which had lasted one month. The results of the first musical memory testing were used as a control measure to equalise high or low initial scores. The difference between the two groups when their adjusted post-test means were compared was 4·32, a result which was found to be statistically insignificant. Gordon had tried during the training periods to offer instruction that would help both the high and low scorers (the children were unaware of the results of the first testing). His results did not confirm the hypothesis that training might be more effective with the more musical subjects. Three of the high scorers and four of the low scorers in the experimental group made sizeable gains on the memory test, but two of the low scoring controls also gained 19 and 29 points respectively. At the end of each week of the training after the second week, Gordon administered a memory test of his own modelled on Drake's. These informal tests gave some indication of growth in progress. Gordon commented that larger groups might have yielded significant differences. We might also wonder whether the training would have been more effective if the low and high scorers could have been taken separately. Longer term instruction might lead to greater improvements. Interest could perhaps be sustained by including the study of themes and variations from published music.

Wing (1948) reported that 25 boys, aged 15–18, improved their initial score by 4·1% when a second testing was preceded by a 20-minute lesson and discussion on the material of one of the subtests. A control group gained only 3·4%. This very small amount of improvement with practice compares quite favourably with research on intelligence tests. Practice on identical tests generally produces a rise of about 5 IQ points, on parallel tests nearly 5 IQ points. Coaching with practice at taking complete tests can produce quite large gains. The total average gain from two practice tests plus a few hours coaching is estimated by Vernon (1960) to be about 9 IQ points for the majority of British town children.

Familiarity with the music of the test items appeared to have little difference on Wing scores (Wing, 1948) even when the proportion of known items was quite large. Most boys from a group of 100 who had considerable experience of music knew only between 5 to 15 out of 80 items, but a few knew up to 45. Yet the percentage of correct answers for known and unknown items was almost exactly the same.

ABSOLUTE PITCH

In Chapter VI we noted that absolute pitch usually develops in early childhood. Is it possible to acquire this ability later in life?

Seashore was convinced the absolute pitch was innate, but several experimenters have succeeded in greatly improving their subjects' ability to name notes correctly. Helen Mull (1925) trained her subjects on the 12 sounds in the same octave. They practised attentive listening, followed by testing, three times a week for four months. At the beginning only 7% of the answers were correct, at the end 62%.

Maltzeva (see Teplov, 1966, p. 178) experimented with five subjects. Three of these were fairly musical; one improved after 12 training sessions from 11% to 25% correct replies, another after 22 sessions from 18% to 36%, and the third after 24 sessions made correct judgments about one-third of the time. The fourth subject, a singer, improved in 25 sessions from 19% to 44%. The fifth subject, a student at the Moscow Conservatoire, improved from 26% to 67% in nine sessions.

Lundin (1963) used a programmed learning technique to teach his subjects to identify a random series of musical tones. Twenty-four chromatic notes from middle C to B^1 were recorded on tape. Each subject worked in a booth. During training he was told immediately when he named a pitch correctly. When he was wrong he was told how far he had erred in number of semitones and in direction from the correct note. He was then required to correct his error by pressing a button corresponding to the right note. During the first part of Lundin's research he trained five male students. Two claimed to have absolute pitch, but this was found to be unjustified by the pre-training test, two of the others had some musical training, the other had had none. After 36 sessions all showed marked improvements. Two of the subjects had gained absolute pitch.

During the second part of the research the notes found to be easiest were presented first. Five subjects worked through at their own rate 14 graded tapes. After ten trials one subject had improved by 700%, another by 600% and a third by 300%. Two subjects who had made only one correct response at the beginning also showed considerable improvement.

Teplov agrees that it is possible to teach with some success a subject without absolute pitch to name the notes of the piano. However, he believes that this artificial ear for absolute pitch is fundamentally different from true absolute pitch in being less precise and more unstable. Far from it being improved by contact with music as is the case with true absolute pitch, it tended to be lost in taking part in music. It seemed to depend largely on timbre whereas true ability to name notes does not. However, unlike Maltzeva's singer, at least one of Lundin's subjects improved his singing on pitch and even complained of increased sensitivity to the faulty intonation of others.

TIME AND RHYTHMIC DISCRIMINATION

As we shall see in the next chapter, rhythmic tests seem to be somewhat more resistant to the effects of music lessons than some other tests. But can performance be improved by specific training?

The evidence is less substantial than that found with pitch, but the answer seems to be 'Yes'. Ross in 1914 experimented with a time test similar to Seashore's repeated seven times on successive days. Fifty trials were given each day at the two time intervals closest to the subject's threshold of the day before. This practice resulted in 'marked improvement', the average gains ranging from 11% to 50%. Klauer of the University of Iowa (see Farnsworth, 1928) gave intermediate grade children two months practice in marching, clapping and beating time. His experimental group showed no significant improvement over the controls, who, however, were retested after a shorter time.

Ashley Coffman (1951) of Northwestern University gave special training to 18 children aged 13 and 14 and to 12 College students. They were selected on a basis of low scores on the Seashore rhythm test and on a rhythm discrimination test of Coffman's own, in which the rhythmic patterns were presented in musical contexts. The 12 training periods, each lasting 50 minutes, were spread over three months. Coffman tried to adapt the training methods to suit the particular needs of the individual. Methods used included clapping and marching to music, drumming and beating time and practice in discrimination. At the end of the training the younger group had improved significantly on both the discrimination tests over a control group who had received no training; the College students had improved significantly on the Seashore test. The experimental group had also taken a motor rhythm test which required them to tap back 60 rhythmic patterns. Both showed marked gains on this task. Five of the College students who played the piano were asked to perform pieces described as 'rhythmic in character' before and after training. On average they were judged to have improved 'slightly' by five trained musicians.

M. T. Henderson (1931) used Seashore's rhythm meter as a training device. After only five days of practice at matching rhythmic patterns on the meter, his nine piano student subjects had made definite progress.

Skinner's teaching machine technique (see p. 158) can also be used for rhythmic training. The subjects' task was to tap in unison with a rhythm presented by the machine. Skinner found that skill could be improved from very approximate agreement to the skilled performance of difficult rhythms. Some of the more difficult patterns taxed even experienced musicians.

Gordon devoted three of his training sessions (see p. 161) to Drake's rhythm test. The control group gained on average slightly more than those who received training. However, individual scores within both groups showed extreme gains and retrogressions. Among the experimental group, seven subjects gained from 1 to 93 points (median 20) and three lost 5, 74 and 89 points respectively. Two of the control group improved their scores and eight showed losses of from 4 to 70 points. Several children made more than the 91 mistakes allowed for in Drake's norms.

It will be noticed that all the above studies refer to activities that would be more properly described under 'time' rather than rhythm in the sense of feeling the onward flow of music. How far the latter can be increased by training has not been investigated.

CONCLUSIONS

Some of the improvements discussed in this chapter are doubtlessly of a purely 'cognitive' nature in the sense that explanation of what is involved in the concept of, say, pitch may help unsophisticated subjects to make somewhat higher scores on Seashore's test.

In addition, there is substantial evidence that specific coaching, if efficiently carried out, can improve specific skills. How far such changes could lead to overall improvement in musical ability is not yet clear from the available evidence.

XVII

The Effects of Music Lessons on Test Performance

As noted in the previous chapter, programmes of special coaching and remedial training usually have to be discontinued after relatively short periods. The only readily available measure of long-term musical training is provided by instrumental lessons. However, only the fact of whether or not the individual has had lessons or the number of years these have continued has usually been taken into account when comparing the amount of training with test results. Qualitative estimates of the instruction have not usually been possible.

In addition to any light which a study of the effects of music lessons on test scores can throw on the theoretical problem of how modifiable is musical ability, it is of some practical importance to the music teacher to know whether or not any allowance need be made for the child who has had music lessons when evaluating test results.

THE SEASHORE MEASURES

Reporting on their experiments during the development of the Seashore tests, Seashore and Mount (1918) reported low correlations (up to ·31) between musical training, as carefully estimated from questionnaires, and pitch discrimination for large groups of students. Only low correlations were found by De Graff (1924, cited Farnsworth, 1928) between the amount of music lessons and rhythm discrimination among several hundred adults and children. With 20 music students, however, Brennan (1926) found significant correlations between the number of half-hour lessons her subjects had had and the pitch test ($r = \cdot 42$) and the memory test ($r = \cdot 55$). Graves's results confirmed the last two of these studies. The experimental group made significantly higher scores than did the children not receiving lessons on the Seashore pitch and memory tests, but not on Seashore's rhythm test (Graves, 1947). She commented that it was difficult to believe that the families sorted themselves so definitely in terms of inborn musical aptitude that only children from such families, who also ranked high in musical aptitude, were given lessons.

Stanton (1922) compared the talent profiles of the subjects who took part in her genetic study (see p. 119) with the amount of their previous musical training. On the following table those classified as grade A had majored in music at college; had spent one or more years of study abroad; had had extensive private study; those in grade C had taken music courses and had several years' experience as players; those in grade E had had no more than three or four years of instrumental lessons early in life, or no musical education at all.

Musical training	Seashore talent profiles Superior	Excellent	Average	Poor	Total
A	9	3	0	0	12
C	8	6	5	0	19
E	8	8	9	5	30
	25	17	14	5	61

While there would appear to be some relationship between musical training and the Seashore profiles, Stanton noted the relatively high proportion of talented persons who had not received musical training.

With her colleague Wilhelmina Koerth, Stanton later (1930, 1933) reported much more extensive data collected at the Eastman School of Music. Four groups of students were tested on entrance to the School and retested three years later. The table below shows the average scores made on the two testings:

Group		Seashore raw scores Test 1	Test 2
I	285 pre-adolescent students in the preparatory dept	76·5	81·2
II	208 adolescent students	81·2	83·3
III	152 special students	80·6	81·9
IV	157 music degree majors	84·0	84·1

According to Stanton and Koerth, the amount and quality of training were similar for the three younger groups and much more extensive and intensive for the adults, yet the mean increases in score were less as development advanced. The improvement among the younger groups was said to be due to the progressive lessening with maturation of 'cognitive' obstacles, and the lack of improvement among the adults was attributed to their having already reached their physiological limit at the time of the first testing. Wyatt (1945), however, pointed out that if students who left the course before the second testing (i.e. the less successful ones had been included) a greater increase might have been found. She pointed out further that 35% of Group I and 20% of Groups II and III had gained 7 points or more, thereby increasing their PR from for example 50 to 90. It is, nevertheless, true that the scores of 57% of the youngest group (i.e. those most likely to improve) had not varied by more than 4 to 6 points. In the case of the third group the small average increase was partly due to the lower retest scores for the highest quartile. The third quartile – where there was more room for improvement – increased their scores by 5·6.

THE K–D TESTS

The average total score of some 4,200 children aged 10 to 19 was 11·25 points higher (out of 275) in the case of those who had received six months or more training (Kwalwasser, 1955). This was particularly true of the Tonal Memory, Tonal Movement, Rhythm Imagery and Pitch Discrimination tests. It is curious that Rhythm Imagery should be more affected than Pitch Imagery. Both tests are likely to be susceptible to training influences since they require a knowledge of notation. When a longer period of instruction was taken as the criterion of the 'trained' group, the mean score increased still further, but Kwalwasser does not mention how the longer period of training affected the subtests.

THE DRAKE MUSICAL APTITUDE TESTS

Drake (1957) claimed that his tests are measures of 'pure aptitude'. The correlation between number of years of music lessons and the rhythm test is certainly very low (·01 with 228 American children and adults and ·26 with 131 Belgians). But the correlations he gives for the memory test

	Score (number of errors)			
PR	Non-musical @ 11 to 12 years	Musical @ 11 to 13 years	Non-musical @ 15 to 16 years	Musical @ 14 to 16 years
25	77	66	70	54
50	70	54	61	44
75	58	44	48	34

and number of years the pupils had had music lessons, though rather low, are by no means negligible: ·37 (with 190 American children); ·35 (with 50 English child music students) and ·43 (with 160 English schoolchildren). In fact, Drake provides separate norms for use with students who have had five or more years musical training.

These differences appear to be quite considerable. They may be partly due to the self-selection of good students tending to continue with their music lessons, while the untalented give up. However, since five years is an arbitrary division, many of the non-musical group must have had music lessons for less than five years and the differences are, therefore, all the more significant. It is, of course, desirable that, where a test is known to be affected by training, definite information should be available, so that the user can make due allowance for previous tuition.

An investigation of the effects of systematic musical training on the development of musical memory was recently completed in Poland by Jan Horbulewicz (1963). The need for such a study arose from the very high percentage – 62% to 89% – of pupils in music schools who failed to complete their courses. This seemed to be due to the adoption after the war, of the equalitarian principle that any individual had unlimited possibilities for the development of any ability.

Horbulewicz used the Drake Memory Test on pupils in music and general schools, aged between 10 and 17, and on groups of professional musicians. As the Drake test was too difficult for a group of seven-year-old children who were applying for admission to an elementary music school, he devised a test of his own. In all, his subjects numbered 473. He concluded that music training was not the determining factor in the development of musical memory. The tempo of development was not quickened by training. Even after more than ten years of a uniform programme of instruction, great differences remained among individuals. The groups that had undergone training did reach a higher level, making about 23% fewer errors on the Drake test, but this was partly due to the wastage among the less talented. The different abilities involved in the Drake test – memory for time, for pitch – seemed to be subject to the influence of training in varying degrees.

THE INDIANA–OREGON MUSIC DISCRIMINATION TESTS

Although Hevner tried to select music which was not likely to be familiar to the ordinary person a correlation as high as ·64 with musical training was found with 126 college students. With advanced musical students, however, scores were not related to training (Hevner, 1931).

Long in his revision also endeavoured to use unfamiliar music. In practice however he is finding that some of his selections from the piano repertoire are well known to piano students. Before his revision of the test is published, he hopes to substitute less familiar music.

THE WING TESTS

To find out whether instrumental training made any difference to the performance of his tests, Wing (1948) divided 271 children into three categories; above average, average and below average. Just as many children having instrumental lessons were in the bottom group as in the top group. The children, however, were only 12 years old; some may have been learning an instrument for too short a time for training to have its full effect. Wing claims that with older children results would be affected by self selection – the weaker pupils giving up lessons and those who are gifted becoming self-taught if no lessons were available. Shuter (1964), in fact, found a correlation of ·65 between Wing scores and a combined measure of their lessons and playing among her sample of grammar school children (see p. 121). She attributed this relatively high correlation largely to the less talented giving up music lessons as grammar school children would be especially liable to have to do because of pressure of school work.

Wing also found that the correlations of the scores of several groups retested after one to five years were high and unaffected by the fact that many had continued studying an instrument, a few had started to learn while others had given up playing. Even in boys' grammar schools where it was usual for all school music lessons to cease when the pupils were 14, ability to do the tests continued to improve till the age of 17.

Wing classified his subjects merely into those who had lessons and those who did not, irrespective of the length of training. Newton (1959), however, classified the RMSM junior musicians into three categories: those who had had three or less, four to six, or seven to fifteen terms of instruction. As they were undergoing a formal system of organised tuition, it was thought that the amount of their musical experience could be directly related to length of time under training. (Data on their musical experience prior to entry were excluded from statistical assessment as being too unreliable.) No significant difference was found between the mean Wing scores of the boys in the three categories.

Holmstrom (1963), however, cast doubt on the value of Newton's evidence. He supposed that the group of junior musicians 'is most likely sampled from the better – from the point of view of environment – half of the population, and has probably a considerable amount of musical training before entering the RMSM'. He argues that on entry to the School, they had already passed the stage where training could affect their scores. However, the proportions of talented, average and unmusical, as reported by Newton, were, in fact, very close to Wing's norms for the general population. As for training, more than half the boys had had less than two years instruction in music, *including* their terms at the RMSM, and only 24 had more than five years musical experience. No doubt, Holmstrom was right to think that the boys were likely to have received some positive encouragement from home or school to join the RMSM and that the proportion who played or listened to music before entry is probably higher than in the general population of adolescent boys. Yet, when Shuter (1964) analysed the Admiralty questionnaires (see p. 149), she found that only 36% gave music as their main spare-time interest and 25% did not play at all in their spare time. (They may, however, have taken part in music-making at school.) 15% did not include listening to music among their spare-time activities. Even if the junior musicians were less sophisticated musically than Holmstrom believed, it might still be true that they had passed the stage where intensive study could improve their performance of the Wing tests.

Whittington (1957) in his original report was inclined to attribute the difference between the musical and non-musical groups to the greater experience of the former. 'Results seemed to indicate that the musical group was superior to the non-musical group because of musical experience, an experience which accounted for some 44% of the performance.' The subjects he tested and later excluded from the non-musical group because they had learned an instrument or had some form of music in their homes made scores a little higher than his completely non-musical group. Later, however, Whittington somewhat modified his views and agreed in 1961 with Wing 'that the relation between musical experience and test results is not that the experience caused the high results but that both spring from high musical intelligence'.

Jack Heller (1962) at the State University of Iowa studied the effect of fifteen weeks of formal music training, where emphasis was placed on the development of listening skills, on the Wing scores of 164 American College freshmen. The training periods lasted 50 minutes and were taken three times a week. The gain, though statistically significant, was small when the experimental group and 41 controls were retested. There were no differences in the improvement on the first three tests compared with the appreciation part of the battery, and none between the group when subdivided into high, average and low scorers.

Holmstrom (1963) compared the grade 2 and grade 4 results of 47 children who had been playing some instrument between grades 2 and 4 with those of 75 children who had not been receiving private lessons. Marked training effects were found for all the subtests used by Holmstrom, especially for pitch. He considered that the improvement might have been greater still if some of the 47 children had not already had lessons before the initial test (Holmstrom stated that he had other evidence which seemed to suggest that training effects on such tests soon reach a maximum). The average improvement was actually not very great: roughly 10 marks, compared with 8 marks made by the control group, out of a possible 104. The effect of the training might, of course, be much greater with some individuals.

SUMMARY

It seems to be true (*a*) that many students with no formal musical training make higher scores on musical ability tests than do some subjects with considerable training, but, (*b*) subjects who have had music lessons do tend to make superior scores.

The question 'Are the superior scores of those who have had music lessons due to their training or to selective factors?' is difficult to answer. Kwalwasser (1955) considers that 'it is much more rational and realistic to maintain that training is a by-product of talent (than that talent is the product of training), for those possessed of talent seek and receive instruction'. Wing explained the higher proportion of musical adults having received lessons in terms of their being free to follow their own interests and to develop their own talents. He suspects that nowadays selective factors would tend to improve the scores even with younger children who are learning music, since parents and teachers are now more conscious of the need to 'spot' and develop talent at an early age. When using tests for prognostic purposes, it seems reasonable to bear in mind that the scores of candidates who have been learning to play may have been somewhat raised by the lessons. But as far as can be judged on present evidence the increase is not likely to be large enough to invalidate the test result.

XVIII

Discussion and Conclusions

It would seem from the previous chapters that musical ability tends to run in families and to appear early in life and in individually varying degrees which do not seem to be consistently related to the amount or quality of environmental stimuli. Such data will be interpreted by many people as supporting the view that musical aptitude is largely innate.

DISCUSSION OF HEREDITARY STUDIES

The data summarised above might be interpreted as supporting the view that musical aptitude is largely innate because:

(*a*) it tends to run in families,
(*b*) it tends to appear early in life and in individually varying degrees which do not seem to be related to the amount or quality of environmental stimuli.

On the other hand, Wallace (1914) contended that musical families like the Bachs and Couperins were merely remarkable instances of the continuity of vocation. Since the child of a professional instrumentalist cannot be brought up in a sound-proof room he hears music from his cradle and later may benefit from tuition from his parent and an easier entrée into the music profession. The child of the composer, painter or poet is less likely to follow in his father's footsteps because the latter works in silence. Lundin (1967), too, considers that studies of family histories can support a view preferring the acquisition of musical behaviour just as well as they can support the inheritance theory. He points to the musical surroundings in which Bach and Mozart grew up and quotes Pronko and Bowles (1951):

> both on his mother's and his father's side for two generations back there was not a single musician in Haydn's ancestry . . . The argument that the musical ability of the Bachs was hereditary because it 'ran in the family' should hold just as consistently for

> their German-speaking activity[1] . . . Their (Haydn's and Bach's) genius behaviour was the culmination of a series of events . . . involving long hours of practice and other labor.

The effectiveness of a mere repetition of stimuli is, however, doubtful. Again, although many composers had good opportunities for the development of their musical powers and were encouraged and well trained in youth, others were either forbidden by their parents to take up music professionally or had little opportunity to do so. Though most composers had to struggle extremely hard before they were able to express themselves effectively, yet, from the consideration of Mozart, Schubert and Berlioz, it appears that the technical as well as the emotional aspects of the creative faculty can, in some way, be partly inborn. Scheinfeld, too, found that some of the greatest virtuosi that he investigated came from 'the humblest and least musical homes; . . . some of the lesser ones from highly musical backgrounds, with both parents professional musicians'. Such a lack of consistent correlation between musical achievement and background would suggest strongly, Scheinfeld concluded that musical talent does not arise from any unusual *home environment*, *per se*. That a highly musical environment also (or alone) cannot produce talent was shown by the children of virtuosi, most of whom showed no unusual talent.

It is not quite true that 'for two generations back there was not a single musician in Haydn's ancestry'. His father had learned to play a harp and, although he could not read music, he delighted in singing alone or with Haydn's mother. Also, Michael Haydn, Joseph's brother, was a church musician of distinction and a third member of their generation had some talent as a singer. Haydn had no children – or none that can be proved his – who might have shown talent.

The forbears of certain musicians who apparently came from unmusical families may have lacked the opportunities to learn music. Had Gluck not been sent to school at the age of 12 his talent might never have developed. The first 12 years of his life were spent in a completely unmusical background, but as soon as he had the opportunity his ability showed itself very quickly. Whether some of his ancestors would have displayed talent given the opportunity is not known. It would appear to be somewhat more convincing to attribute Gluck's ability to an untraceable hereditary factor than to an environment known to be unstimulating musically.

[1] A more accurate analogy would be between speaking German and writing contrapuntal music.

Mjoen (1926) acknowledged a category of individuals whose talent could not be explained by the ability displayed by their parents or collateral relatives. To explain the biological appearance of eminent ability, Mjoen considered it was not sufficient to work with average values and quantitative investigations, because the nature of a quality might change under the influence of other qualities. Genius might be explicable in terms of combinations of congenital talents.

Lundin (1967, p. 205) states that he has 'no objection to the concept of inheritance, per se, providing we try to discover what it is that we inherit', and further (p. 222) that what he is arguing against is 'the inheritance of mental powers for musical reception and performance'. In his view musical 'capacity includes, among other things, a sound nervous system, two hands, normal hearing structures, and other structures necessary for musical behaviour' (p. 206).

The writer would prefer to formulate some tentative answer to the question 'What is inherited?' in rather different terms. One answer might be 'such genes (or more likely combination of genes) which predispose the individual to perceive, remember and judge music more (or less) efficiently than others not so endowed'. In Chapter XX the importance of the higher mental processes in musical perception will be argued. However, it seems possible that the 'higher mental processes' complex though they are, may be ultimately explicable in neurological and biochemical terms. If Lundin's 'sound nervous system' is meant to extend to the higher brain centres involved in musical perception and judgment, it might be tentatively accepted as adequate on this account. More important is the question how adequately does Lundin's definition deal with the wide range of individual differences which are evident from the norms of music ability tests and other studies of the musical activities of young children. Does Lundin mean *any* child with a sound nervous system and normal hearing structures can develop a high degree of musical capacity if brought up in a musical background? Lundin agrees that deficient structures will obviously be limiting factors for musical behaviour and that the man born deaf is deprived of part of his biological equipment with which he may acquire musical responses. But it is not clear whether, for example, tone deafness would be considered due to a 'structural deficiency'. Lundin does recognise that 'all people with similar training do not achieve the same degree of proficiency in musical tasks' and that 'these limitations are a function of both biological capacity and previous musical experiences'. Lundin thus seems to have moved away from the extreme behaviourism of J. B. Watson, who claimed that he could 'guarantee', given a free hand in controlling a child's environment and training, to take any normal infant 'and train him to become any type of specialist I might select – doctor, lawyer, artist, merchant-chief and, yes, even beggar-man and thief, regardless of his talents, penchants, tendencies, abilities, vocations and race of his ancestors'. Lundin is willing to consider that the rate of maturation, a function of heredity, is important. It is easy to agree that musically precocious children often have a head start because of early development, so long as he does not mean to imply that the average child could 'catch up'. The more outstanding at least of the musical prodigies continue to develop far beyond the heights of achievement attainable by the ordinary person even with considerable training. Mere rate of maturation does not provide an adequate explanation of their exceptional degree of talent.

The Soviet psychologist, Leontiev, took the view that individual differences in special abilities like the power to reproduce sounds of a given pitch correctly only seem to be inherited. This is because 'the reflex systems of which they are a function are formed only under certain definite conditions they do not, therefore, always develop and, in the case of different individuals, they may have a different structure'. But he does not seem to provide a real explanation of why some children fail to form the 'links' and connections essential to the building up of the mechanisms involved in for example the correct vocalisation of musical sounds. If this depended mainly on the environment, it would be reasonable to expect a clear relationship between the individual's ability and the amount and quality of his musical experience. Admittedly, the environment even for siblings is never identical. However, the available evidence on within-family differences, and on differences between individual families and the population in general, support the importance of innate factors in setting an upper limit to achievement and speed of learning.

There is ample evidence that efficient programmes of aural training can improve the performance of certain prerequisite skills such as pitch discrimination and singing. How far-reaching the effects of such gains could become under optimum conditions is not known. In some cases perhaps all that would be achieved would be the attainment of the ceiling of development at a rather earlier age.

PART IV

Theories of Musical Ability

XIX

The Nature of Musical Abilities

In the preceding parts of this book we have seen how psychologists have tried to devise measures of musical ability, how it seems to develop from the earliest years of life to adulthood, and how far heredity and environment play a part.

But what *is* musical ability? Is there in fact a unitary ability, or should we rather be speaking of musical *abilities*? Is it more realistic and useful to consider musical ability as one broad factor, different aspects of which may be sampled by different tests, or as composed of separate abilities which, however, may overlap and seem to work together? Is there some distinction between sensory and 'higher' levels of musical ability? If so, how are these related? How important is musical memory? Can some abilities be regarded as 'basic' in the sense of being preconditions of the development of others?

In trying to answer such questions, the psychologist considers how far the various tests intercorrelate – and may go a stage further and by statistical analysis seek to discover whether performance of different types of test depends on a few underlying factors. A summary of factorial studies can be found in Appendix II.

THEORIES OF MUSICAL ABILITY

While the nature of musical ability is generally admitted to be complex, opinions differ on the extent to which its various aspects intercorrelate consistently.

At one extreme, Seashore believed that musical capacity may be divided into a number of sharply defined talents which are unrelated and can be present or absent in varying degrees. His tests correspond to the physical properties of sound: pitch, time, intensity and timbre. He claimed that such capacities are as basic to musical aptitude as they are to sound itself. Moreover, 'each one of these capacities runs as an independent branch, not only in sensation, but through memory, imagination, thought, feeling and action' (Seashore, 1938). Thus, a 'sense of rhythm' depends on the basic capacities of time and intensity discrimination, and memory upon pitch. Mainwaring and Arnold Bentley agree with Seashore in believing that musical ability is composed of separate elements. Bentley, for example, states that all the functions measured by his tests, though they may overlap and usually seem to be working together are in fact separate (Bentley, 1966). If this view is correct we should expect to find that intercorrelations among music tests are very low and that the search for underlying factors would yield only highly specific ones.

On the other hand, Wing believes that there is a general ability to perceive and appreciate music. Strong intercorrelations exist between tests designed to measure the different aspects of musical ability and an important general factor underlying all such tests will be found. In addition to the main factor, he recognises other minor factors: one sorting the tests and the persons tested into two main types – analytic and synthetic respectively, and one dividing ability to judge or perceive harmony from ability with melody or rhythm. Wing did not deny the possibility that a separate factor of rhythm might exist, though he obtained no evidence of it with his own test material. Both McLeish (1950) and Shuter (1964) support Wing's view on the essential oneness of musical ability.

An in-between position has been adopted by others such as Holmstrom (1963) who argues for a number of group factors. From his Uppsala study, Holmstrom claimed to have obtained three musical ability factors: an Alpha factor, based mainly on pitch, a Beta factor dependent chiefly on factors of experience and memory, and a Gamma test performance factor related to intellectual level.

Another theory of some importance is the Integrative theory of musical ability, associated with Drake (1933). While adhering closely to the Seashore theory of specific talents, he thought that they might all depend on, or be knit together by, musical memory.

Let us look at each of these theories in the light of empirical evidence, bearing in mind that the results of correlation studies and the factorial analyses derived from them depend on the tests chosen and the subjects tested. The reliability and validity of the tests are also important in assessing the meaningfulness of the results.

THE THEORY OF SPECIFIC CAPACITIES

Correlation studies of the Seashore measures do suggest that they test relatively distinct abilities. Yet zero correlations are rare. The following table shows the range of intercorrelations which Teplov (p. 63) tabulated from 14 studies, and the average correlations which he computed:

TABLE XIX.I

	Range		Average
Pitch–intensity	·09 to	·82	·33
Pitch–time	·11	·54	·32
Intensity–time	·08	·95	·33
Pitch–memory	·30	·77	·52
Pitch–consonance	·13	·78	·34
Intensity–rhythm	·07	·46	·22
Time–rhythm	·00	·55	·32

Eight of the correlations of pitch with memory exceeded ·5, but the connection between rhythm and time or intensity was no greater than the intercorrelations of the supposedly 'unrelated' capacities.

Some of agreement between the correlations is probably attributable to intelligence. But, even with intelligence held constant, Drake found a common factor and two group factors, and McLeish a general factor. In one of Drake's group factors, intensity was linked with pitch, and in another, with time. He concluded that even when a special attempt is made to measure isolated and independent abilities, it is seldom absolutely successful. However, he agreed that there was more that was specific than common in what the tests were measuring. In McLeish's study, this view was confirmed – factors specific to each test accounted for 38% of the variance compared with the 24% of the general factor. In Holmstrom's re-analyses of studies including the Seashore tests, most of the test variables proved to be loaded in more than one factor, and several factors had significant loadings in more than one of the Seashore subtests (Holmstrom, 1963, p. 88). The theoretical division of auditory functions according to the physical features of the sound wave (frequency, intensity, complexity and duration) did not correspond to the factors found in Karlin's 1942 study, nor to Holmstrom's re-analysis confined to Seashore tests. Pitch was closely connected with quality (timbre) and intensity with time.

Mainwaring, like Seashore, took as the basis for his tests the four physical attributes of sound, though he confined his tests to pitch and time, and to recall. The correlations he obtained were indeed low (range ·14 to ·36). This is hardly surprising considering the nature of the tests (see p. 279). Fieldhouse (1937) even found negative correlations between the Mainwaring pitch and rhythm tests, though, before rotation, the rhythm test did have a small loading on a general music factor.

With Arnold Bentley's tests, all his intercorrelations were in the 40s, except between pitch and rhythmic memory and between tonal and rhythmic memory. In a factorial study carried out by a Polish student, Kinga Donat, the Bentley tests, applied to 59 secondary modern English schoolchildren, did form part of a general music factor (after rotation). But, we may agree that at the level of these tests intended for younger children and where for the sake of simplicity time and pitch elements have been separated, memory for pitch does seem to be fairly distinct from memory for time patterns.

If pitch and time constitute 'independent branches' throughout music, we might expect Lundin's melodic tests to be but little related to his test of rhythmic sequences and the same to occur in the case of the Gordon tonal and rhythmic subtests. The relationship between the Lundin melodic transposition and sequences tests, and the rhythmic sequences test is indeed quite small (range of correlations ·31 to ·39). As mentioned in Chapter II, the validity of the Lundin rhythmic test also seems to be quite low. It is perhaps not without significance that the music teachers whose assessments were used as the validation criteria were not willing to try to rate rhythmic behaviour separately.

A much closer connection between the tonal and rhythmic parts of Gordon tests was found both by Gordon himself (1965) and by Tarrell (1965). His own figures were based on the results from all the pupils in grades 5, 8 and 11 whose scores were used in the standardisation of his tests and range from ·45 to ·65 for the subtests and from ·59 to ·71 for the combined tonal score compared with the composite rhythmic score. With nearly 1,500 children in every grade from 4 to 12, Tarrell's intercorrelations ranged from ·35 to ·70 for the subtests, and from ·55 to ·73 for the composite scores. In spite of these substantial correlations, Gordon points out that the values are lower than the reliabilities of the subtests and claims that this 'constitutes evidence although somewhat indirect, of the multi-dimensional nature of musical aptitude'. In an earlier section of his test manual, however, he shows that he is well aware that as they exist in music, 'rhythm and melody interact in an inseparable way'. Gordon tried to obtain from music teachers evaluations of specific musical abilities which would correspond to his seven subtests. In most cases the specific ability, e.g. for singing, in parts did correlate rather more highly with the subtest in question, in this case the harmony part of the tonal tests, than with any of the other six tests. But all the correlations were low.

THE THEORY OF A GENERAL MUSICAL ABILITY

When considering Wing's claim that there is a general factor common to all musical activities, or at any rate common to the activities involved in the music tests, we must bear in mind that, in order to produce an economical battery of tests, he eliminated from his final seven certain tests which had high correlations with some of these, and which were therefore regarded as redundant.

The range of correlations obtained with his tests vary considerably with the group studied. With the unselected schoolboys whose scores were used for his main 1941 factorial study, more than one quarter of the correlations were in the 50s. McLeish, too, found quite high correlations among the chord analysis, pitch, memory and harmony tests. On the other hand, Shuter's homogeneous groups yielded many low, zero or even negative correlations. These studies, however, invariably provided evidence of a strong and important general factor, accounting for up to 40% of the variance. (All factor analyses of music ability tests do, before rotation, show a general factor, but some authors do not publish the amount of the variance.) Which tests contribute most to the general factor seems to depend partly on the composition of the group. The phrasing test appeared to be the most efficient with Wing's group. With College students of average musical ability, both McLeish and Shuter found that memory and pitch were the most important tests. The chord analysis test assumed greater weight in the case of Shuter's two highly musical groups.

How far any common factor found by factorization is evidence of the unitariness of musical ability depends, of course, on how successfully the tests used covered all aspects of the ability. The comprehensiveness of the Wing battery depends on how wide was his original choice of tests and how valid is his claim that 'no vital test is missing from the short series' (Wing, 1948, p. 49). The 7 tests of the short series correlated very highly with 13 wider ones. These in turn gave very high correlations with the original 24 – as comprehensive a battery of music tests as has ever been used.

Mainwaring (1947), however, expressed some doubt as to how well the short series covered the rhythmic aspect of music. He agreed that, if we could regard a general factor of music as having been established, then the most valuable tests would be those most saturated with it. But, on the other hand, if, as Mainwaring himself believed, musical ability consisted of a group of independent variables, then the weak association between a test and the total of a battery would be all the greater reason to *include*, rather than to *exclude* it. Wing had, however, found that a test of time-pattern dictation from his longer series and rhythm appreciation (as included in the short series) were the two tests which were among those with the 12 highest intercorrelations (out of 171 intercorrelations of 18 tests). It therefore seemed reasonable to consider one redundant. The rhythm appreciation test was chosen for inclusion as the rhythm test of the short series because firstly 'the tasks set can be made extremely difficult without needing, or being noticeably influenced by, any acquired knowledge of musical technicalities. Secondly, it is closer to the appreciation required in listening to normal music. Finally, it can be applied repeatedly to the same people without deterioration in efficiency' (Wing, 1941*a*, p. 143). In his review of the Wing tests, Mainwaring (1948) did concede that as 'a preferential judgment implies a preliminary perception of difference and, in this kind of judgment, the ability to differentiate between pattern and distortion, it seems reasonable to assume that the inclusion of one rhythm test . . . is in this instance quite adequate'.

With the possible exception of the recently published Gordon tests of tempo and metre, other types of time and rhythm tests do not seem to have been proved particularly useful or valid. The two Gordon tests appear to be reliable and of promising validity, but each takes 18 minutes to administer. Certainly, if only one test could be included, Wing's would seem the most likely to be valuable, as the Rhythm is presented in a musical context. It would, of course, be interesting to study whether giving the Gordon tests in addition to the Wing battery would increase its predictive value.

A test which Wing himself was forced to omit 'with regret', owing to the length of time it took to play was appreciation of tempo. A reviewer of Wing's tests (Buros, 1953) suggested that this item might have been shortened rather than discarded. However, even with only two choices of tempo instead of three, or with the number of items reduced from 20 to 14, it would have still required nearly 30 minutes to administer. (What Gordon calls his test of 'sensitivity–style', must be rather similar, since the only difference between the two renditions is a change of tempo. In any case, the person who can make a high score on all four of the Wing appreciation tests, would probably be perfectly capable of choosing a suitable tempo for performance (within his technical capabilities, but that is another matter).)

It might be argued that the Wing battery lacks a test of pitch discrimination for finer intervals than the semitone. The Wing pitch change test, however, has been found to correlate quite highly with

Seashore's measure of pitch (r ranges from ·48 in Franklin's smaller study to ·74 in McLeish's results). The Wing had a greater weight on McLeish's general music factor when both batteries were considered together. The Wing test was more discriminating than a 35-itemed version of the Seashore in differentiating between Faulds' specialist music group and the Princeton students. In practice, there is of course no reason for not giving the Seashore pitch test as well as the Wing in situations where a fine degree of pitch discrimination is important.

GROUP FACTOR THEORIES

While some psychologists have been content to interpret their results in terms of one major factor of general musical ability and several minor ones, others have preferred to try to identify a number of 'group factors' of more nearly equal status. The results of trying to carve up musical ability have not always been too satisfactory, since clear and consistent divisions are rarely found. The statistical procedure used is the rotation of axes (see Appendix II). But, as Wing pointed out (1941*a*, p. 279), 'with the present tests this rotation is an extremely hazy business, for we cannot . . . assume that any one of the given tests is a pure test of a particular function in music'.

As mentioned above, Karlin found that the physical properties of the sound wave could not be relied upon to suggest divisions that were psychologically meaningful. In his earlier studies where several tests of musical relevance were used, he claimed to have found evidence for a pitch factor and for two different types of memory factor. As no less than three memory tests out of eight or nine auditory or musical tests were included, it is not surprising that some distinction between 'memory for form' (i.e. the ability involved in Drake's musical memory test) and 'memory for elements' (as in Drake's retentivity test) emerged. If a sufficient number of tests concerned with one particular aspect of musical ability is given, unless the tests are very similar, some differentiation is likely to be found. This is analogous to inspecting a leaf under a microscope in order to distinguish the different type of cells.

Franklin was the first to present a group factor solution in studies including the Wing tests, the Seashore pitch and memory tests, as well as his own TMT test and two rhythm tests adapted from Revesz. Franklin, however, did not make an issue of whether there was one musical ability or several. His main interest was arguing for a distinction between 'mechanical-acoustic' pitch discrimination and 'judicious-musical' pitch discrimination. By mechanical-acoustic discrimination

Franklin meant the ability to discriminate between very minute pitch differences as in the Seashore test. Judicious pitch, as required by the Wing pitch test, was the ability to deal with pitch changes in a musical context. He claimed that there was a good measure of agreement between his own results and those of Wing and of McLeish. The common ground between his views and Wing's is the belief in the inadequacy of sensory discrimination tests. It is McLeish's remark that the Seashore tests measure the same kind of ability as Wing's, but at a different level that causes Franklin to feel that there is a measure of agreement between his own and McLeish's results. Franklin, however, argues that judicious pitch is the most important aspect of any general music factor that may be found.

Holmstrom, the most sophisticated of the factor analysts who have been concerned with musical ability, pointed out that the individual with fine pitch discrimination is likely to be helped rather than hindered when he attempts the more musical tests. By re-analysing Karlin's, McLeish's, and Franklin's figures, using a more objective rotation procedure. Holmstrom believed that he had uncovered two group factors 'in which the loaded variables in both cases are concerned with pitch discrimination'. In addition in one of these factors elements of intensity discrimination occurred, in the other, elements of tonal memory. The first factor he called 'Alpha'. This was a primary pitch perception factor, which had possibly a physiological basis and might be only slightly influenced by experience of music. It differed from Franklin's mechanical-acoustic pitch factor in not being restricted to the perception of minute differences. He believed that on the psychophysical level of the Seashore tests, an intimate connection exists between variations in frequency and intensity, making the intensity loadings on the Alpha factor plausible. The second factor, Beta, was the one in which the more musical tonal memory tests had their greatest loadings. This distinction between pitch-intensity and memory-pitch factors does not always seem valid – a fair loading on memory appeared in the Alpha factor in his re-analysis of Franklin's earlier study. Pitch and intensity did not appear together in one of the Karlin re-analyses. But, Holmstrom holds, we must not interpret our factors only in terms of test content, but also with reference to differences in the amount of experience of music to which the people we are studying have been exposed. Pitch, he believes, can be affected by experience, thus with older subjects with varied experience of music, pitch enters increasingly into the Beta factor, relatively little remaining in the Alpha factor.

In his own studies, Holmstrom found further evidence of Alpha and

Beta factors, and also of a third factor, which he called Gamma. This Gamma factor appeared among the Uppsala children at grade two. It was a broad musical factor, loaded particularly on rhythm, pitch and memory, with strong connections with intelligence. It was no longer present in the results from grade four. He interprets it as a general test performance factor, representing ability and ambition to do well in a testing situation. With the 'B' group from unmusical homes, knowledge of music was related to this factor and not to the general scholastic factor, as was the case with children from normal or musical homes. A higher general intelligence was thus, Holmstrom concludes, of little help in answering questions about music in the groups whose environment had been musically poor. That such a Gamma factor can be demonstrated among children of eight or nine does seem quite plausible (cf. p. 227 below).

In grade two an Alpha factor appeared which persisted into grade four, though its loading on memory virtually disappeared and its loading on pitch became much lighter. By grade four a Beta factor emerged due, according to Holmstrom, to the greater experience of the older children with music. With the group from unmusical homes, the Alpha factor at grade two was associated with attitude to music and to perception of pitch, as well as marks for music. A Beta factor again appears at grade four, after school music has with time had its effect of making up for this group's poor home background.

Holmstrom's factorial solutions are undeniably elegant and legitimate, but how true are they to reality? Are there alternative interpretations? It is certainly possible that musical ability *may* be usefully divided into those functions particularly dependent on perception of pitch and another concerned with memory, i.e. with the laying down of engrams in the cerebral cortex. The factors thus delineated are not clean cut; we should not expect them to be. What is open to doubt is whether the latter may not be just as 'physiologically' or genetically based as the former. Alpha functions may require a genetic basis and less experience than Beta functions, but the evidence reviewed in Part Three seems to make clear that *both* are partly dependent on genetic control. Again, perception of a melody as a whole seems to be present from a very early age.

What alternative solutions might be proposed to Holmstrom's data? If we look for a general music factor, there will be no difficulty with his Enköping study, as there is only one musical ability factor. With the Uppsala grade two analyses there are two. If we consider the Gamma factor as being the one representing general musical ability, we should have to say that chord analysis has not yet developed sufficiently in these younger children (cf. Chapter VI) to be part of their musical ability. The Alpha factor would seem to be especially concerned with pitch, although memory has, with the 'B' group, actually the higher loading. The Beta factor is probably closest to Wing's general music factor (as indeed Holmstrom himself remarks). In his re-analysis of Franklin's larger study, leaving out the repetition of the Seashore pitch test, the first factor could be considered a general one. It is in fact quite similar to that of McLeish and that of Shuter with comparable subjects.

THE INTEGRATIVE THEORY

The importance of memory in musical ability is undeniable. An appreciation of form can hardly exist unless the listener can recognise themes when they return at a later point in the composition. Through memory the performer can store up models of excellence of tone quality, of phrasing or of dynamic expression. Again, in Drake's words, 'Memory functions primarily to make it possible for the rendition of a piece to have unity, meaning, variety and individuality. To interpret a composition intelligibly it is necessary to perceive the piece as a whole, as well as the relationship of all the parts to each other' (1939).

Some examples of empirical evidence on the importance of memory may be quoted:

1. Of all the tests used by Bentley (1955) the Wing memory test gave the highest correlation with the total score. This is not surprising since many of the other tests depended on memory – the number of tests involving memory in the field of musical ability is in itself some indication of the importance attached to memory by test authors.

2. The high discriminatory power of the Gaston melodic memory test in Bentley's study; and the good validity of Drake's memory test and of Seashore's tonal memory test.

3. Fieldhouse's finding that a defective auditory memory was a more important cause of singing out of tune than faulty pitch discrimination. Fry (1948) reported results which tentatively suggested a similar conclusion. However, a group of monotones tested by Arnold Bentley were significantly below normal in both tonal memory and pitch.

Gordon (1965) seems to underestimate the value of musical memory as a function to be tested. At least when discussing the rationale of his own tests he lays down the principle that a high degree of musical memory would not be necessary for satisfactory performance of his tests since the use of repeated melodies and rhythms reduces the amount of new material that the students can listen to during a given span of time. However, memory certainly enters into the performance of all his tests. In fact, the instructions explicitly ask 'remember the song . . .'

Important though memory may be, does it in fact function, as Drake has suggested, to knit together discrimination of pitch, time and rhythm, etc.? One way to study this question is to see whether the intercorrelations between such tests drop substantially when memory is held constant. Bugg and Herpel (1946) investigated this point. The following table shows the original correlations between the various Seashore tests when they were given to 181 subjects and the correlation that remained when memory had been partialled out statistically.

TABLE XIX.2

Variables	Original *r*	Partial *r*	Diff.
Pitch–Timbre	·45	·31	·14
Pitch–Rhythm	·34	·21	·13
Pitch–Loudness	·32	·20	·12
Rhythm–Timbre	·22	·12	·10
Pitch–Time	·22	·13	·09
Rhythm–Loudness	·37	·31	·06
Rhythm–Time	·35	·31	·04
Loudness–Time	·43	·40	·03
Loudness–Timbre	·00	−·11	·11
Timbre–Time	·03	−·05	·08

Holding memory constant does, in fact, reduce the intercorrelations but except in the cases of pitch–timbre, pitch–rhythm and pitch–loudness, not very greatly. Bugg and Herpel point out that memory correlated ·60 with pitch and ·30 with rhythm and that we might, therefore, expect that a considerable amount of the overlapping between these tests would be due to memory. They neglect the possibility that what is common to these tests might, for example, be the willingness to concentrate on auditory material and need not necessarily be memory. A more general discussion on the place of memory in musical ability will be found in the next chapter.

RHYTHM AND MUSICAL ABILITY

A separate factor of rhythm seems to lurk in the shadows of several studies where tests of pitch and memory have been highlighted. Both Karlin and Wing acknowledge the possibility of its existence.

Inspection of many of the correlation studies shows that tests of rhythm seem to be the odd man out (McLeish (1950); Bentley (1955); Rainbow (1965)). In both McLeish's and Wing's factor analyses, the rhythm tests had the lowest loading in the general musical ability factor. In Franklin's study, the two tests that he adapted from Revesz form a separate factor of their own, unrelated to the Wing rhythm test. From some evidence, it would appear to have strong connections outside music (see Chapter XXIV).

One of the difficulties of studying rhythm is that subsumed under the heading of time or rhythm are a motley collection of tests whose reliabilities in many cases are poor and whose relevance to music is, to say the least, dubious. Intercorrelations between such tests as the Seashore measures of time and rhythm, and Drake's rhythm test, and Lundin's are very low, as are those between Wing's and Revesz–Franklin's, and the other rhythm tests included by Bentley in his study. Yet, given a test like Wing's where a musical context is used, rhythm has an appreciable (if relatively lower) loading on the general music factor. As already mentioned, the Gordon tests of Tempo and Metre have moderate connections with his tonal tests.

An interesting study of rhythmic aptitudes has recently been reported by Hiriartborde (1964). His subjects were 65 young women who were training to be teachers of physical education – a rather select group. Besides the Wing and Seashore test batteries he gave patterns of short and long notes, and of soft and loud stimuli to be reproduced and written down, and a variety of motor rhythm tests. He extracted five factors. Before rotation there was a general factor with loadings on all the tests included in his analysis. After rotation he obtained two 'musical' factors, one had high loadings on the Wing Intensity and Phrasing tests and on the Wing and Seashore Memory tests as well as a loading on a test of synchronisation. The second seemed largely concerned with pitch and time and showed little connection with his non-musical tests. The factor which showed the highest variance, 28%, was a complex one involving motor co-ordination and control of movement, besides tonal memory and pitch. The second most important factor with 23% of the variance was interpreted as perceptual ability to structure rhythmic groups. The only music test to have an appreciable loading was Seashore's rhythm test. His last factor was concerned with ability to synchronise responses with rhythmic stimuli. Emotional stability as measured by Cattell's PF. test had appreciable loading on this factor. This seemed understandable as nervousness on the part of the subject would handicap him in synchronising his movements with the stimuli.

Although Hiriartborde administered the whole Cattell battery, only factor 'C' showed any relationship to his music and rhythmic tests.

Thackray (1966) reported some preliminary results that he had obtained by factorizing his aural and visual perception tests. Although all the tests proved to be highly specific, in both cases he found small general factors. The tests of rhythmic patterns, counting, and comparison of rhythms had the highest loadings on both the visual and the aural general factors.

Interesting though these findings are, how closely Thackray's and Hiriartborde's tests are related to functional and musical ability has still to be established. Olin Smith (1957) showed that the Seashore rhythm test is quite closely connected to following a pattern on a rhythm meter and also continuing to tap after the clicks had stopped.

The connection between rhythm as it is involved in music and as it exists outside music is, of course, an interesting area for intensive study. Such investigation might throw light on many problems within the field of musical ability. However, as Teplov points out, rhythm in music is qualitatively different from rhythm in general. Musical children found Revesz's rhythm test easier when it was presented in the form of tunes, unmusical children when the time patterns were tapped. Rhythm in music is not only connected with motor movements, but with the emotions expressed by the music. An appreciation of rhythm should be connected with dynamic changes, with harmonic progression and with phrasing, as well as with perception of time and pitch differences.

ABSOLUTE PITCH AND MUSICAL ABILITY

How is absolute pitch related to musical ability? Stumpf went as far as to say that absolute pitch was necessary in order to understand and fully enjoy the great masterpieces of music. Others have pointed out that it can be a nuisance, because it interferes with activities such as transposing.

Absolute pitch is sometimes found in unmusical persons, but this seems to be rather exceptional. Many of the great composers possessed this ability. In a series of experiments carried out on 18 very talented children at the Moscow National Conservatoire, Blagonadejina, found that 13 certainly had absolute pitch (Teplov, 1966, p. 195). Sergeant (1967) found a similar proportion among the most talented of his four groups.

The advantages seem to outweigh the disadvantages. Absolute pitch is an aid in developing an ear for harmony as well as allowing modulations to be followed with ease. Individuals with absolute pitch are not, however, necessarily conscious of the modulations when listening or performing music, according to Weinert (1929). Weinert questioned 10 experienced musicians on this point; he also asked whether being constantly aware of the modulations was a decisive factor in the pleasure they obtained from music. Only two agreed that it was. Sergeant (1967) asked 46 musicians, 23 possessing absolute pitch and 23 without, to listen to the opening 15–20 bars from 20 familiar compositions. Ten had been transposed, ten were played in the original key. Care was taken to select music in which the opening theme did not occur in any other key elsewhere in the work. The musicians were asked whether each excerpt was played at the pitch at which they were accustomed to hear it. The mean score of the absolute pitch group was slightly lower than that for the non-absolute pitch group. Sergeant concluded that the pitch level at which a piece of music is performed does not play a more important part for an absolute pitch subject, and that the attention of the mature musician is centred on other aspects of the music.

This does not mean, however, that absolute pitch is not helpful in the development of musical ability. Since it facilitates a more analytical perception, it may be used as a means to the end of understanding of the music.

Teplov would seem to sum up the consensus of opinion when he concludes that absolute pitch is very favourable to the development of an ear for music, but that it is by no means indispensable.

INTEREST IN MUSIC

To be talented is not necessarily to be interested, nor to be interested, talented. Every music teacher has pupils of very considerable ability who 'could not care less' and others of moderate aptitude who show a keen interest and desire to learn. Yet, on the whole, interest in music and musical ability are likely to be associated.

Wing (1948, p. 78) compared the grades of 333 adolescent boys on his test with their own assessment of their degree of interest in music. The coefficient he obtained was approximately ·30. More recently, Leon Crickmore has compared the Wing scores of two groups of students attending a Liberal Studies course at a technical college with their liking or disliking of 30 gramophone records. The pieces ranged from popular to classical music. The longest record lasted about seven minutes, the shortest about two. The seven point scale on which the students rated each piece ranged from 'disliked very much' (1 point) to 'liked very much' (7 points). The total score probably gave a reasonable measure of their attitude to music in general. The correlation coefficients between the liking-scale and Wing were ·27 and ·31. Similar figures were obtained by Lundin between his tests and liking for classical music.

Both Shuter (1964) and Rainbow (1965) worked on the basis that children who were interested in music would spend time listening to 'good' music. Shuter compared concert-going and home listening with the Wing scores of the grammar school children she tested (see p. 121). A very low correlation resulted (·1). She believed that this was partly due to pressure of schoolwork on the children, competing interests, and to the more gifted preferring to spend their leisure time making music rather than listening. In assessing the interest in music of his three groups of schoolchildren (see p. 146), Rainbow awarded points for reading about music and musicians, and for the probable purchase of records (given the money). To check the type or types of music the child claimed to like he was asked to name a composer, performer or work associated with each type. Rainbow counted jazz as equal to serious music for the purpose of evaluating music listening habits. The correlations he obtained between interest in music and the Seashore pitch, memory and rhythm tests, and Drake's memory test were reasonably close to Wing's figure of ·30. The highest correlations were with the two memory tests and ranged from ·28 to ·42. The lowest were with the rhythm test: ·17 to ·27. A significant correlation of around ·40 was found between interest in music and the teachers' ratings which he used as a criterion of musical ability.

Holmstrom (see Appendix II) assessed separately his subjects' attitude to music as a school subject, and the attitude to music that they had acquired at home. For the latter they were asked to compare a hypothetical programme of music with five non-musical programmes. Holmstrom gives no details of the non-musical programmes. If one or two were very attractive, many children who enjoyed music might have selected them in preference to music. On the other hand, if unattractive programmes were included, children who did not particularly care for music might have chosen music rather than a programme they liked even less.

Both measures of attitude showed little relationship to the music ability tests except with his group from unmusical homes. In his smaller investigation both were related to school music marks and to some degree to knowledge of music. This was less apparent in his Uppsala study where he obtained a greater number of factors. However, the attitude to school music in grade 2 did seem to bear some relation to music marks in grade 6 and might be of prognostic value. This result agreed with the experience of music teachers that attitude to music becomes increasingly important for success in school music as the child grows older. Lack of interest is a very frequently mentioned reason for giving up music lessons (Casey, 1964; Martignetti, 1965). With the children from musically poor homes, attitude to music was strongly linked with music marks *and* with an Alpha musical ability factor.

We may conclude that, at any given level of musical talent, interest in music is likely to be an important determining factor in whether or not the child's potential capacity is fully realised. Speaking from his experience of having trained over 3,000 young pianists, Cortot (1935) concluded that an unshakable devotion to the art was even more necessary than a natural flair. Whether or not the young artist has this devotion to music cannot be proved till he has encountered the tremendous obstacles that beset his path. We may compare this conclusion with Terman's finding that at the highest level of talent (as shown by his gifted group) interest and personality factors are crucial to success in life.

XX

Perceiving, Remembering and Judging Music

The purpose of this chapter is to examine more closely what we mean by the ability to perceive, remember and judge music.

In order to perceive clearly, we must first be prepared to attend. Secondly, we must be able to hear the music; so that we may enquire therefore whether or not auditory acuity has any effect on musical ability. Not only must we be able to hear, we must be able to discriminate differences of pitch, loudness, timbre and time. Even more, we must be able to grasp the music as a whole and do so, not at our own speed, but at the tempo dictated by the composer. Ability to relate tones so that they have meaning and represent a musical idea would hardly be possible without memory for music. Beyond memory a judgment level may be reached where we can, for example, compare two renderings of a composition and decide which has the more artistic worth. Finally, we must not overlook the contribution which our muscular system makes not only to *performance* but also to *perception*.

ATTENTION

A primary requirement for musical perception is that we should be attending to the music.

Perhaps the first point to be made about attention is to note how limited are our powers to attend to the world around us and within us. Indeed one of the great difficulties for the inexperienced listener is that so much is going on at once. No sooner has he grasped the subject of a Bach fugue, when a second voice starts up; he is just beginning to cope with the two melodies, when a third voice enters, followed by a fourth or even a fifth. All evidently have something important to say, the original subject is recognised from time to time, but while he is attending to that, he is missing something else. Perhaps, he decides, it would be better to give up and just enjoy the flow of the music. This leaves out of account extraneous stimuli – our neighbour's cough, her low-cut dress, or the hardness of the seats. Listening to music with sympathetic companions and watching an instrumentalist or orchestra may often be a considerable aid to the apprehension of the music (Vernon, 1934). Sight and sound stimuli seem to travel by separate channels up to the level of the brain where they have to be sorted out and co-ordinated; again, listening to music in a not too dimly lit concert hall may help to keep us from feeling sleepy as the evening draws on. However, it is all too easy to be distracted from the music, especially when, as with a televised concert, our visual stimuli are chosen for us.

Because our powers of attention are limited, we tend to select from our environment those stimuli which seem most worth perceiving. Some of the research on selective listening carried out by psychologists and others interested in communications would seem relevant to problems of listening to music. Cherry (1953), presented his subject with two speakers, one to each ear on headphones, and asked him to listen to only one of them. To make sure that he kept his attention fixed, Cherry asked him to repeat back the words to himself as he heard them, keeping two or three words behind, but talking continuously. The subject was later asked what if anything he had noticed about the speech to which he was not attending. The answer was very little indeed; he would be aware, for example, that it was a woman speaking on his left side, but he would fail to notice that she switched from English to German. That it was possible to notice some general properties of the unattended speech without being aware of its verbal content showed that the perception of speech occurs in at least two successive stages and the limit to our capacity for attention arises chiefly at the second stage, at which we identify what is being said. How efficient may be the screening mechanism which enables us to be selective in our attention was neatly demonstrated by R. Hernandez-Peon at the University of California. He implanted electrodes low in the sensory pathway of a cat's ear. This enabled him to record the nerve impulses produced by the sound of a metronome. Just as soon as a more interesting stimulus, such as a mouse or a goldfish was introduced into the situation, the response to the clicks was greatly reduced (Morgan, 1965). Again, we may not notice the clock ticking till it stops.

As we saw in the last chapter, interest and ability do not always go hand in hand. Mere willingness to attend is no guarantee of effective perception. No doubt a habitual tendency to learn through listening or what Drake (1940) called 'earmindedness' (a possible explanation of his common factor of musical ability) may be connected with musical aptitude. The musical person is likely to find more to interest him in music, because he can perceive details that sustain interest and is more sensitive to the emotional expressiveness of the musical progression. But it seems best to distinguish interest from ability.

ACUITY OF HEARING AND MUSICAL ABILITY

Acuity of hearing does not seem to be related to performance of music tests.

The possibility that defective hearing might be one cause of singing out of tune was investigated by Fieldhouse (1937). His 50 subjects were schoolchildren aged 9 to 11. They had been selected by their class teachers as being unable to sing in tune, had scored five or less at some simple ear tests and had sung two songs of their own choosing off pitch. Fieldhouse tested them with the Seashore battery and with Mainwaring's tests of pitch and rhythm, as well as with an audiometer, and compared the results with the performance of 96 'normal' children. The greatest difference in average score between the tune deaf and the normal children occurred on tonal memory, but the differences on the pitch tests were also significant. The auditory acuity test gave zero or even slightly negative correlations with the pitch and memory tests for both groups.

Wing (1948, p. 67) tried out his tests on 22 boys who were considered sufficiently deaf to need medical treatment. He concluded that deafness up to a loss of 15 decibels, the limit of the figures used, did not affect ability to do his tests.

Tomatis (1953), however, claimed that the musical ear is characterised by a progressive rise in acuity between 500 and 2,000 cps. While investigating the hearing of some 1,300 aircraft workers, he found that a certain proportion, 4%, appeared to have suffered no hearing loss even after being exposed to levels up to 130–140 db for over 20 years. Their right ears, on the contrary, showed improved hearing in the crucial 500–2,000 cps region. On the other hand, two of his patients, professional singers who complained of having lost the power of singing in tune had audiograms similar to workers whose hearing had been damaged by noise. Tomatis concluded that the singers had been deafened by the power of their own voices. In so far as he succeeded in improving the acuity of their hearing in the region of 2,000 cps their capacity to sing in tune was restored. Tomatis then enquired into the musical ability of the 50 workers who seemed invulnerable to noise. Without exception, all were musicians or at least could sing back a phrase in tune. Thus, he claimed, the audiogram could provide a quick estimate of whether or not a person was musical.

It is reasonable to argue that what a person can't hear he won't be able to reproduce. However, granted that all the 50 workers were musical, 4% sounds a very small percentage of musical ability even in a select population. We are left wondering whether some of the workers whose hearing had been damaged by exposure to noise might not also have been musical.

Farnsworth (1941) obtained evidence that the more musical of a group of college students had better acuities especially between 510 and 2,000 cps but he did not give separate values for the right and left ears. He used two criteria of musical ability: self ratings on a five point scale and Seashore test scores. In an earlier investigation with two groups of over 1,000 schoolchildren, the only significant difference between the musical and unmusical was for the right ear at 1,900 cps. In this case the *unmusical* group had the better auditory acuity (Farnsworth, 1938).

SENSORY DISCRIMINATION

How important for musical perception is the ability to distinguish fine differences of pitch, time, intensity and timbre is a question that has been argued ever since Seashore claimed that his tests were fundamental to musical ability. Leaving out of consideration tonal memory which he believed was less basic than the sensory capacities, only his pitch test seems to have much proven validity, as we saw in Chapter II.

Precisely how keen pitch discrimination needs to be for practical musicmaking has not yet been established (Bentley, 1966). One difficulty is that persons with equally fine powers of sensory discrimination do not necessarily apply them in musical situations to the same degree. It certainly seems reasonable to agree with Holmstrom that ability to differentiate small pitch intervals is more likely to help than to hinder the perception of musical tasks. The correlation between the Wing and Seashore tests of pitch is usually found to be relatively high (average correlation from three studies exceeds ·5). The Wing test is related to functional musical perception in requiring detection of a change of pitch against the masking effect of the other notes of the two chords. It differs from melodic tests in involving only very short-term memory.

A more useful division than between micro intervals and larger than semitone ones, might be between tests involving intervals of two chords and those concerned with melodic pattern. In R. R. Bentley's study, the Gaston test where the subject is asked to find a given note in a chord was only slightly related to his test of phrases moving up or down; Whistler and Thorpe's test of melodic recognition was not closely related to their test of pitch discrimination, though fairly highly to their pitch recognition test. The relationship of the Wing pitch test to Whistler and Thorpe's melodic recognition test as well as to their pitch tests was moderately close. But Lundin found that his interval test did not correlate highly with his tests of melodic transposition and sequences. Petzold found that the ability to learn a musical phrase is not strongly influenced by the accuracy with which children can sing back short melodic items. High scores on his 45-item test (see p. 82) did not insure that phrases incorporating these same items would be learned. Effects of memory are introduced with increasing length, as well as clearness of perception.

Franklin refers to physiological evidence for a distinction between simple pitch discrimination and the perception of a succession of tones. This was an early experiment reported by Pavlov. Three years after the removal of certain higher brain centres concerned with hearing, a dog could be trained to discriminate between sounds, but was unable to distinguish an upward moving scale from a downward moving one (a task well within the powers of a normal dog). More recent research with animals has shown that the highest part of the brain (the cerebral cortex) is not needed for discrimination between sounds differing in pitch or intensity. Removal of the relevant parts of the cortex produces complete destruction of ability to perceive the difference between two patterns, e.g. ABA as contrasted with BAB (Neff, 1964).

How this finding applies to humans is, however, uncertain. In the case of sight, destruction by disease or accident of the parts of the cortex concerned with visual perception results in a complete loss of sight and not just perception of pattern. The physiological mechanism for the perception of pitch is not fully known, but the orderly projection of the sensitive area of the inner ear on to the higher brain levels is probably connected with the resolution of pitch; for we find that stimulation of specific spots on the membrane of the inner ear seems to be projected to corresponding spots in the auditory area of the cortex. The inner ear (cochlea) itself can analyse sounds and separate one tone from another, if the frequencies are not too close together (von Bekesy, 1957). In the higher nerve centres the rough analysis is made sharper. The 'sharpening' of the differential effects of acoustic frequency seems to depend on a particular frequency being not only able to *excite* certain nerve cells of the auditory system more than others but also to *inhibit* certain cell units. Thus what starts out in the inner ear as the excitation of a relatively large group of receptor cells can be narrowed down to a smaller group higher up by the inhibition of units that also respond otherwise to other frequencies (Morgan, 1965). The relative efficiency of such an inhibitory mechanism in different people could produce individual differences in keenness of pitch discrimination.

From the evidence discussed in Chapter XVII, it would seem that practice can often improve pitch discrimination. However, Seashore is right that there must be an ultimate physiological limit to fineness of pitch discrimination. The breaking of the 'mile-in-four-minutes' barrier in running should lead us to be cautious in postulating what might be a limit to human pitch discrimination. It is interesting to note that athletes compete for their medals after periods of very stringent training.

But would an Olympic feat of pitch discrimination in itself make a musician? Wing (1948, p. 12) cites the case of a boy whose percentile rank on the Seashore pitch test was 100, but who could not sing the simplest song in tune, although the quality of his voice was good. He had had lessons on the violin for a considerable period, but had given up in despair as he could not learn to play in tune. Why should this be so? The answer lies in the importance of being able to relate the separate pitches into a melody, to perceive them at speed and to remember them, in any actual musical (as opposed to laboratory) situation.

PERCEPTION

Mursell, Vernon, Wing and Revesz, to name but a few, have all stressed the importance of being able to perceive a tune as a whole rather than as a succession of notes. Revesz (1953), for example, says 'The melody consists of single notes, just as a figure is made up of lines. However, the final product that is directly perceptible is a functioning whole, as individuality that cannot be apprehended from the parts . . . The melodic-rhythmic impression of a musical motive remains the same' even when the tune is transposed, while, conversely, the change of a single note may change the entire character of the melody. He goes on to state that the totally unmusical person views a melody as a mere sum total of notes and fails to perceive it as a whole pattern. It would, however, seem more likely that most unmusical persons hear a tune as a whole, though they fail to perceive the relations of the notes to each other and to the tonic. In fact many naïve listeners react to musical language as young babies react to words or as adults react to a foreign language. They cannot identify words, nor reproduce them, nor apprehend the syntax or logical relations. They can only tell whether the tone of voice is pleasant or unsympathetic, and their reactions are largely emotional. But as the infant grows older or the adult becomes accustomed to the strange tongue, they gradually differentiate the verbal noises, and learn to recognise simple sentences and later more complex speech (Vernon, 1934). Sensory discrimination enters into the perception of the melody since we have to be able to distinguish differences of pitch, intensity, time and so forth. But in music these differences are not isolated as in a laboratory but exist together. We must also be able to generalize, for instance, when we recognise the function of cadences in musical phrases.

Not surprisingly Spohn (1965) with his programmed learning experiments (see p. 261) found that whereas a time-pattern was easy to learn and an interval moderately easy, when the two were combined to form a tone group, the task became much more difficult.

It seems, however, that it is not the sheer number of notes that may make perception difficult. At least from her experiments at the Medical Research Council Psycholinguistics Research Unit at Oxford, Anne Triesmann (1966) has concluded that the limiting factor is not the number of sensory stimuli but the rate at which the brain has to resolve uncertainties about their nature or identity. If an analogy with speech can be drawn, this would mean that it is not the total number of notes that is important in musical perception but the complexity of the ways in which they are classified and analysed. With practice, we learn to interpret stimuli more quickly.

An important factor governing the rate at which the brain can perceive is expectation. The more unexpected the stimulus, the more questions the brain must ask to determine its identity. As we might expect, strings of unpredictable words, are much more difficult to repeat back than passages of coherent prose (cf. Kate Gordon's experiment on p. 203). The development of a system of expectations (of a return to the keynote, of which chord is likely to follow, etc.) is of considerable importance in the perception of music. The development of an ear for melody seems to depend largely on the establishment of a sense of tonality. The skilful composer to keep the listener alert will, of course, often deliberately introduce the unexpected. Unless the listener knows what the expected would have been, he will not fully appreciate that what he is hearing is unusual. Much of the fundamental emotional expressiveness of music is derived from the arousing of expectations which are – ultimately – fulfilled.

Another aspect of being able to perceive a whole from reduced cues is the ability of the advanced musician to 'fill in' in imagination a suitable harmonisation when he hears a melody.

Tom Ritchie (1960) at the University of Indiana tested the hypothesis that an accompaniment of simple diatonic chords would aid the perception of four-note melodic fragments. Given the first note, the subject had to write the other three notes of the melody. His subjects found that an unharmonised version was easier than the harmonised one. When secondary triads were used, perception of the melody became more difficult still.

MEMORY

Unless we have the memory for music of an infant prodigy, or idiot savant, we are not likely to be able to remember an unknown tune after only one hearing. The sensory impression fades rapidly and must in some way be consolidated in order to be retained. Problems of memory have been extensively studied by psychologists using verbal material; memory for music would seem to be analogous.

One obvious factor that contributes to the difficulty of memorising is length. Three words or three notes are much more easily reproducible than ten or twenty. As the number of notes is increased the proportion remembered drops, although the absolute number will increase. The span of immediate memory, the number of digits or letters that can be repeated perfectly after one hearing, is limited. About six or seven is the usual span for digits, for letters of the alphabet it is about five. However, if the letters are combined into words, then the span becomes five words. Experiment has shown that it is not the amount of information that determines the span of memory but the number of units or 'chunks'. Since the memory span is a fixed number of 'chunks', we can increase the amount of information that it can contain by building larger and larger 'chunks'. Seventeen unrelated letters would be almost impossible to remember. But if the letters happen to be 'THECATSATONTHEMAT' as soon as we recognise how the unit is organised, memorising it becomes easy.

Even musicians may have difficulty in identifying isolated intervals, especially the tritone, minor sixth and seventh. Teplov found that music teachers would argue about an interval, an argument that would only be settled when one of them quoted a passage of music in which the interval occurred. Again, James Marquis (1963) at Iowa University found that ability to perceive the basic quality of intervals in melodic sight-singing is considerably less important than ability to perceive the scale, harmonic and tonal changes surrounding the intervals.

If we have to try to identify objects that differ from one another with respect to a single aspect, our capacity to do so is surprisingly limited

(Miller, 1964). For example, a subject can reliably distinguish among about four different loudnesses when the intensity is varied, or among about five different pitches, when the frequency is varied. If both aspects are varied at once, we might expect to find that the four loudnesses could still be distinguished at each of the five pitches. In fact, however, only 10 or 12 tones can be recognised with both aspects varied. Persons with absolute pitch who can identify accurately any one of 50 or 60 different pitches are what Miller elsewhere (1956) called 'remarkable exceptions' to the general rule of limited capacity for making absolute judgments of auditory stimuli.

In the jargon of communication theory the process of building larger and larger 'chunks' is known as recoding. The most usual type of recoding in everyday life is a verbal one. If we want to remember a story, we re-phrase it in our own words and remember the words. Verbal recoding may be useful in trying to remember music. We might recode the opening of the *Leonora* Overture and remember it as a descending scale. Recoding becomes easier if we can make use of redundancy, i.e. if we know which letters or notes are likely to follow which. In English for example, 'c' is much more likely to be followed by 'a' than by 'b' or 'd'. Miller and Selfridge (1950) re-arranged words in different orders according to the chances of their being found together in English syntax. The closer the sequences of letters (or words) were to real words (or sentences), the easier they were to remember.

Some musical sequences are easier to remember than others – those that conform to our expectations are easier to identify and to recall. 'Nonsense' music is difficult to learn because it does not conform to our expectations.

Kate Gordon (1917) played five melodies backwards to 20 adults who had been classified as musical or unmusical on the basis of their performance of reproducing the melodies in their original form. As we might expect, the musical group were more affected by the change. None the less, they were much more efficient at singing the melodies played backwards than were the unmusical group at reproducing the tunes when played forwards. In time, the 'nonsense' music came to sound like tunes and performance of both groups improved.

The primary auditory image seems to retain the characteristics of the particular perception. Whipple (1901) experimented with single tones. He found that the primary auditory images weakened after two seconds, though, with practice, they could become clearer and last longer. With melodies, Blagonadejina (see Teplov, p. 288) found that immediate auditory images could be evoked without conscious effort and that they retained the characteristics of the perception, including the timbre and absolute pitch of the notes. Even a subject who denied having auditory imagery could evoke a primary image during the first few seconds after hearing the melody. He could also sing back the tune perfectly, if it was not too long. This could hardly be due to muscular memory, because he never sang. Blagonadejina believed that this subject retained the auditory image of the first note and that when he sang the first note, its sound evoked the second and so on.

But prolonging the image of a melody immediately after it has been heard is far easier than being able to imagine it or sing it later without the aid of the external stimulus, and is only a stage in the development of the true 'inner ear'. The low correlation between tests of auditory digit span and musical ability tests is indicative of the primary auditory image not being highly important in music. Faulds' group test of auditory digit span produced zero correlations with his other music and auditory tests. Fry's Number Memory test showed little connection with memory for music. The test consisted of a series of numbers presented in pairs through a loudspeaker. The subject had to write down the number that had been changed the second time, e.g. 18749632 as compared with 18759632.

For the development of long-term memory some generalisation has to occur, so that we recognise the *Londonderry Air* whether it is sung by a soprano or a baritone, or played by a violin or the chimes of an ice-cream van (cf. Pflederer's distinction between memory and conservation, p. 206). In the intermediate stages of developing an 'inner ear', recognising a tune is easier than singing it or playing it by ear. Again, we may be able to recognise when someone else is singing or playing out of tune without being able to sing or play in tune ourselves.

Though we may remember a tune in an 'abstract' form, the memory traces produced by each hearing may also be stored in the brain. At least that seems to be the conclusion suggested by Wilder Penfield's demonstration that weak electrical stimulation of the surface of the brain, in the temporal lobes, can sometimes cause a previous experience of the person to intrude into his field of awareness. These vivid flashbacks of previous experiences occurred during the course of operating on patients with temporal lobe epilepsy at the Montreal Neurological Institute. In order to distinguish healthy from unhealthy brain tissue, electrical stimulation was applied to different areas, the patient being fully conscious and able to report his experiences. One woman heard an orchestra playing a tune while the electrode was held in place. The music stopped when the electrode was removed and returned when it was reapplied. She could hum the tune, which always seemed to begin at the same place and to progress at normal tempo. So vivid was the experience that the patient was convinced that a gramophone was being turned on in the operating theatre on each occasion. The patients seemed to hear a single playing of the music by a particular pianist on a particular occasion and might be aware of themselves as being present in the room and feeling whatever emotions they had felt at the time.

The storage in the brain of memory traces will be further discussed in the next chapter. It is not clear from Penfield's report whether the patient would have been able to sing the tune on request under normal circumstances. Presumably these detailed, vivid memories are not ordinarily accessible to recall. The real problem of memory may be one of retrieval rather than of storage. One of the differences between the auditory imagery of the musician and that of the unmusical person is not that the individual without talent lacks auditory imagery. On the contrary, he may quite often find tunes 'running through his head'. However, unlike the musician, he cannot recall at will any particular tune. Musicians themselves differ in the clarity of their auditory images. Even some composers find it helpful to work at a piano.

The connection between perception and memory must be very close – the more definite and 'meaningful' what is perceived, the more likely it is to be remembered; the better past musical experiences have been remembered, the more readily familiar elements in a new pattern of notes can be fitted into existing schemata. If perception is faulty, immediate and subsequent attempts to reproduce a tune may be faulty. Thus, some of Burroughs and Morris's (1962) subjects made errors in their first attempt to sing back a 12-note tune. In later trials, these errors tended to be perpetuated. It was not what the child heard but what he thought he heard which stuck. Immediate memory is important in building up stores of memory traces which will help identification of future percepts. Because of the rapid fading of the primary image, pitch discrimination normally becomes more difficult if the second tone is sounded after an interval of even six seconds (Faulds, 1959). The effect of the delay, however, can be reduced, if rehearsal of the tone is allowed to keep alive the image, or if the subject can make some response that identifies the tone (Bergan, 1966).

JUDGMENT

Some musical tasks require us not only to be able to discriminate between two items, or to remember them long enough to compare them, but also to reach some kind of judgment about them, e.g. on which has the more artistic merit.

An example of a musical function that seems to require something in addition to memory is the ability to recognise a melody when it is played on a different instrument or at a different tempo. Pflederer (1966) is seeking in her current research to separate this 'conservation' function from memory. If we play a tonal pattern to a child followed by versions in which the pattern is 'disguised' and ask him whether the tonal pattern itself has been changed, incorrect responses may be due to failure of either conservation or memory. But, if we include some items that are exact repetitions of the pattern, the errors on these would be indicative of failure to remember. If these items were all answered correctly, we might conclude that memory was playing little part. The extent to which introducing altered stimuli produced more errors would indicate the degree to which the pattern could be conserved.

Drake (1931) concluded from his factorial study that 'judgment, an entirely different operation from memory, is yet impossible without the latter'. Only exceptionally was success on the K–D tonal movement test not accompanied by a good score on tonal memory. However, many did well on tonal memory but not on tonal movement. Bugg and Herpel (1946) found that Seashore's tonal memory test correlated quite highly with the Oregon test (·65) and with the K–D tonal movement test (·61). Again, subjects who scored highly on the Oregon and tonal movement tests did well on tonal memory. But the converse was not true.

Having to decide which of two versions is the better one is a much more discriminating task than merely having to state whether the second version is the same or different from the first. Kyme (1956) found that when he scored his Esthetic Judgment test as a simple discrimination task the scores bore no relation to teachers' estimates of the students' ability. But when he scored it as a test requiring the organisation of the elements into meaningful wholes, the mean correlation between the test scores and the teachers' ratings rose to ·74. His subjects were students taking part in the University of California Demonstration School. With less experienced subjects, the results might have been different.

As mentioned in Chapter VII, the ability necessary to obtain better than chance scores on the Wing 'appreciation' tests seems to develop later than that required for the ear acuity tests. With groups of average ability, memory and pitch tend to be the most important components of a general musical ability factor. At the level of the professional students of the Eastman School of Music, however, Shuter (1964) found that the appreciation tests had gained in prominence. Even when the factors are rotated, these tests do not seem to form a separate 'appreciation' factor. Only in a minor factor did Wing find the tests of rhythm, intensity and phrasing contrasted with the first three tests. From questioning persons who had performed the tests, Wing found that many tended to perform these three appreciation tests by 'intuition' rather than by adopting an analytical approach.

The writer has not yet seen any factorial studies of the Gordon tests. However, correlations between his 'Sensitivity' tests and the rest of the battery seem to increase between grade five and grade eleven from an average of ·32 to one of ·46.

We may conclude that so-called appreciation tests seem largely to measure how well the more primitive perceptual and memory abilities have been used. Probably without a basis of good aural ability, this type of judgment would never develop however much music was heard.

MOTOR PERFORMANCE AND FEEDBACK

From the motor areas of the cortex, neurons descend down through the spinal cord, and connect with motor neurons which extend to the muscles. There are many descending pathways from the cortex, but the area making the largest single contribution is the so-called motor area which occupies part of the frontal lobe. Electrical stimulation of the motor cortex in humans and animals can evoke movements in various muscles of the body, thus enabling the motor cortex to be 'mapped'. It is interesting to note that the amount of motor cortex which controls a certain part of the body is related not to the actual size of that part of the body, but to the complexity of movements of which that part is capable. Thus the areas controlling the hands, and the lips and the tongue are relatively great.

From the muscles sensory messages stream back to the brain, providing continuous feedback. A dramatic demonstration of the importance of feedback in the control of speaking and singing is to introduce a time delay by means of a microphone, and receivers over the speaker's ears. If the delay is sufficiently long, even the most loquacious speaker can be made to falter and to stop speaking. As we noted on p. 157, Tomatis was able to correct singing faults by introducing corrective feedback. Edwin Foot (1965) at Kansas has shown that reduced or enhanced loudness feedback can cause changes in the intensity of singing. He made the point that damage may be done to the voices of young singers by prolonged practice in the acoustically treated rooms so often provided as practice rooms in music schools.

Not all impulses to the muscles result in an audible sound or in a visible movement. Singing to ourselves or feeling the music in our fingers help to keep alive the fugitive memory traces till they can be consolidated. Vernon (1931, p. 126) emphasised the importance of this kinaesthetic aspect in making our perception of music definite. Thus, he states 'as soon as the listener becomes consciously aware of temporal (or pitch) motions in the music itself, external to himself, the response seems to become primarily (though by no means wholly) a kinaesthetic or muscular one'. He quoted a study by Bingham, who concluded from measurements of certain muscular contractions in listening, that attention to simple melodies always involves minute adjustments in the body. Mainwaring (1933) also carried out a demonstration of the importance of kinaesthetic factors in the recall of musical experience. He asked an experienced singer to avoid singing a hymn tune to herself while it was being played to her. After hearing it played seven times, she could sing only the first, second and last bars. Two days later her performance from memory was similar. When allowed to sing to herself during the playing of a second hymn tune, she learned it quickly and was able to recall it with accuracy. More recently in Russia Il'ina (1959) has tried to investigate to what degree motor vocal reactions participate in the formation of auditory concepts in the young child. At least in the case of her 30 musically backward pre-school children, she concluded that they were of considerable importance.

Vernon believed that there were probably considerable individual differences between listeners in the extent to which they depend on this muscular form of perception. Some highly trained musicians may be able to listen actively in purely auditory and intellectual terms. Vernon tried a simple experiment in which he asked two people to imagine to themselves a familiar tune. At the same time he either played a record of some other music, or else asked them to repeat subvocally 'ah, ay, ee, ay, ah . . .' The first subject was relatively unmusical; when his vocal cords were simultaneously engaged he was quite unable to continue imagining the tune. But he was not at all disturbed by the external music. The other was a good musician with well-developed auditory imagery. His imagining of the tune was unimpaired by the activity of his vocal cords, but was completely obstructed by the music from the gramophone record.

Motor factors are most likely to be important when the memorising or recall requires some definite effort (see further Teplov, p. 318).

To sum up, though 'stages' or 'levels' in the perception and integration of music in the brain can be distinguished, each part seems to be intimately interconnected with the others; any divisions are thus arbitrary.

XXI

Physiological Correlates of Musical Ability

Neuropsychologists believe that a valuable contribution to our understanding of psychological functions may be obtained by seeking physiological correlates in the brain. A. R. Luria, Head of the Moscow Institute of Neuropsychology, quotes the example of a famous Russian composer who suffered injury in the temporal lobe of his brain. He lost the ability to discriminate 'p' from 'b' and 's' from 'z'. Yet during the three years that his aphasia lasted, he wrote his best symphonies. This suggests that speech and music are separate functions. Orientation in space, ability to count and ability to understand complicated grammatical sentences appear to be unconnected; yet, if damage occurs to a certain region of the brain, all three abilities may be lost. Luria believes that there may thus be some one factor underlying these seemingly unconnected functions.

Eventually the neuropsychologists may provide the equivalent of a factor analysis of human abilities. Results obtained with brain injured persons should be compared with findings from experiments with normal people. Doreen Kimura did this in her bilaterality experiment described below.

AMUSIA

When we learn something, our nervous system is somehow changed. This change is called the 'engram' or memory trace. The question 'Where in the brain are the memory traces stored?' has long been debated. After exhaustive researches with animals, Lashley (1950) concluded that it was not possible to demonstrate specific localisation of the memory trace anywhere in the nervous system. Penfield's results mentioned in the last chapter seem to contradict this view. However, the memories evoked might not necessarily have been stored in the temporal lobe, since exciting the temporal lobe may excite circuits of neurons extending into many different parts of the brain.

Some evidence for a strict localisation has been put by Karl Kleist (1962). A neurosurgeon during the First World War, he later tried to apply his knowledge of the effects of localised brain injuries to establishing an anatomical basis for sensory aphasia and amusia. After fracturing his skull, one patient, whose case he describes, was still able to tune his zither very accurately and play a few isolated chords. But he could not play melodies, even with the music in front of him. The part of his brain required for comprehension of tones (the first transverse gyrus) was intact, while the subdivisions required for the comprehension and performance of melody and for the comprehension of the text and appreciation of melody had been damaged. A woman who could no longer recognise melodies and had difficulty in copying someone else singing was nevertheless able to sing a song when shown the title in writing, thus proving that the memory traces were intact. Her impaired perception of pitch seemed to be due to subcortical damage involving the transverse gyri.

Trying to connect the loss of specific musical functions with localised brain injuries is potentially a means of understanding the physiological correlates of musical ability. One difficulty is that the damage may have widespread effects. Another is that in the past accurate records of the loss in terms of musical capacities have not always been made. Therefore, Botez and Wertheim (1959) of the I. P. Pavlov Neurological Institute in Bucharest have sought to devise a series of 45 tests which would be systematic enough to be useful and at the same time adaptable to differences in the musical and mental level of the patient. They describe the results of applying the tests to two patients. The first was a farmhand with no musical training, but who had made a living playing the accordion in a band. A year after an operation to remove a frontal tumour, he complained that he could no longer earn his living by his music. He could discriminate semitones, sing back a tune and state the number of notes in a chord. He reproduced piano notes a fourth higher or lower. Though he knew what was wrong, he was unable to correct it. When asked to sing an unknown tune, he could get the general shape of the melody, but his time values were completely wrong. He could tap back a time pattern if played on a piano, but not if sounded with a hammer. He could not play his accordion with both hands together. He seemed to have lost the ability to abstract sounds. He seemed to find more difficulty with intonation on an ascending than on a descending scale, just as children do.

Their second patient (Wertheim and Botez, 1961) had been a symphony orchestra violinist before suffering a sudden stroke at the age of 40. When examined by them a few months later, his amusia was predominantly receptive (i.e. on the perceptual side). He could reproduce single notes and hum a Brahms waltz, but made frequent pitch errors in Boccherini's *Menuet célèbre*. Singing an unknown melody was impossible, as was the minor scale. Major scales could be sung correctly only if he was given the first note. Though he had previously had absolute pitch, he could name only two out of seven notes correctly. The note he named was a fourth above (cf. the accordionist). The main theme from the first part of the *Eroica*, which he had so often played, was now remembered only if played on the piano. In trying to recognise tunes, he would attempt to play them with his left hand (his right being paralysed) on the violin, like the aphasic patient who had to move his lips to understand what he was saying. The rhythm of a melody dictated to him was well perceived, but not the pitch. He could tap back a simple time pattern, but made mistakes at naming duple or triple time. He could still read the treble clef, but had difficulty with the bass clef and found the alto clef impossible.

This accords with the usual finding that earlier learning is better preserved in cases of brain injury than later acquisitions.

MUSICOGENIC EPILEPSY

Some 40 cases in which epileptic seizures have been precipitated by various kinds of music or auditory stimuli have been reported in the literature (see Poskanzer, Brown and Miller, 1962).

In most cases the emotional response of the patient to memories stimulated by the music seems to have been an important factor. The seizures of a man of 24 originated from the undulating note of air-raid sirens. Later he suffered seizures caused by vibrating noises and by long notes on the trumpet or saxophone. A 54-year-old man sought relief from loneliness and emotional disturbances by listening to classical music. He immersed himself so completely that within a few months he knew in detail all of Beethoven's symphonies. Before long he found that while listening to these symphonies his consciousness would be clouded and he would feel an immense joy. Two years later such a feeling was followed by a convulsion and for the next 17 years he suffered grand mal attacks precipitated by music (see Daly and Barry, 1957).

In the case reported by Poskanzer, Brown and Miller themselves, there seemed to be no emotional response associated with the sound. Their patient was a Civil Servant of 62 who otherwise enjoyed excellent health and was an emotionally well-adjusted personality. Six years previously this man had lost consciousness on three occasions when listening to the radio at 8.59 p.m. He later realised that on each occasion a peal of church bells was being played as an interval signal. He had one attack in his car during a programme of church bells on the radio. Eight years prior to the onset of the attacks he had struck his head above the left ear and lost consciousness for a brief period. Poskanzer, Brown and Miller were able to carry out experimental studies with the co-operation of the patient in which they used a variety of sounds one after another at intervals until a seizure occurred. Church bells playing between the frequencies of 290 and 1,120 cps and often in a much smaller part of the band could elicit attacks. The bell-like impact of the sound was essential to produce a seizure; no seizures occurred without the stimulus of church bells. The patient himself felt quite certain that no emotional association was being evoked.

At the moment the studying of cases of amusia and of music-induced epilepsy may not seem to be contributing much that is useful to our knowledge of musical ability. But, with the accumulation of accurate reports, information of importance may some day emerge.

BILATERALITY AND MUSIC

The primary projection areas for auditory messages reaching the brain seem to be located in the temporal (side) lobes of the cortex, which is divided into two symmetrical hemispheres.

The division into hemispheres raises the questions: Do we need both sides or only one? Do we normally use both sides? If we lose the use of one as in cases of a tumour or of an epileptic focus, can the other take over its job? The right hemisphere is concerned with impulses coming from or going to the left side of the body, the left hemisphere with those affecting the right side of the body. Most of us have a 'major' side and a 'minor' one – we are right (or left) handed. Many functions, e.g. speech, are, however, bilateral – we have only one set of vocal organs. Again, in looking or listening we use both eyes or both ears. But even these psychological functions tend to depend more on one side than the other. The left side of the brain seems to be commonly dominant in speech, the right in non-verbal intellectual functions such as music.

Brenda Milner (1962) of the Montreal Neurological Institute of McGill University obtained evidence of the greater importance of the right side of the brain for music by studying 38 patients who were undergoing an operation in which one temporal lobe had to be removed for the relief of focal epilepsy. Twenty-two patients were having the left lobe removed, and 16 the right. Milner gave the Seashore tests to 27 patients before the operation and then two weeks later; 11 other subjects could be tested after the operation only. The groups were well matched with respect both to age and intelligence. Removal of the left lobe had little effect on the tests. The tests most affected by removal of the right lobe were the timbre and memory tests, though all except the rhythm test showed some increase in errors after operation. Milner did not obtain any follow-up evidence as to whether the remaining left lobe was able in time to take over the musical functions of the right.

A further demonstration using actual music and normal healthy subjects was carried out by Doreen Kimura (1964) also at the Montreal Neurological Institute. Twenty female nurses, all right-handed, served as her subjects. She selected 80 excerpts from solo passages in concertos by Mozart, Telemann, Bach, etc. These were classified into four sets of 20 so that within each set the same instrument was used and the pitch range and tempo were very similar, so that the main clue was melodic pattern. The original passages were then re-recorded to make melodies lasting four seconds. For each set two of the four melodies were first played to the subject on two separate channels, so that one melody was heard in each ear. After four seconds silence, the four melodies were played one after the other in normal binaural fashion. The nurses had to identify which two of the melodies had first been heard dichotically. Performance with the left ear (right side of the brain) was again found to be significantly better than with the right ear. A similar, but much smaller, difference was found when clicks were used instead of melodies. The asymmetries were found to occur only under conditions of trying to hear two melodies at once. In an unpublished study Kimura gave the Seashore timbre test to normal subjects, one ear at a time, and found no difference. One reason may be that dichotic listening makes more demand on the system than does monaural.

The research mentioned above seems to contradict Tomatis's claim (see p. 157) that in right-handed people it is the right ear that has the 'musical' audiogram and that lack of a dominant right ear may result in tone deafness and speech defects such as stuttering. However, Tomatis's work in the field of music seems to have been mostly with singers. Perhaps *sung* words like spoken words depend on the dominant hemisphere.

Since speech and verbal ability seem to depend more on the left than on the right side of the brain, Delacato, the director of a reading clinic in Philadelphia, argues that reading disabilities are due to the failure of the left side of the brain to become dominant. To establish dominance, he advocates that during remedial training all singing and listening to music should be cut out. Among the cases he describes was an 11-year-old girl who displayed a considerable gift for music, was ambidextrous in her piano playing, but poor at reading and writing. On Delacato's advice, she was forbidden all music for six weeks and only allowed to practise with her right hand on a silent keyboard. At the end of this time, her left-hand playing had deteriorated and her absolute pitch judgments had become uncertain. After two months of remedial reading her language abilities had greatly improved and she had become very happy and successful at her school work. Her parents, however, were unhappy about the (apparently permanent) deterioration in her musical abilities. We may wonder whether the girl herself had been unduly forced to concentrate on her piano playing and was glad of the excuse to devote less time to music. While music teachers might not object to music being cut out of a child's timetable for a short period, they could hardly be expected to accept Delacato's recommendation that musical activities should be reduced to a minimum between the ages of 5 and 6 as a preventive measure without convincing proof that such activities really did endanger the development of verbal abilities. After all, many individuals are about equally good (or bad) at music and language. Those interested in music education might wonder whether Delacato's prescription could be reversed in the case of children who were verbally adequate but musically deficient!

XXII

Creating, Performing and Listening to Music

CREATIVE ABILITY

As Moya Tyson (1966) remarked, there seems to be a growing interest in the study of creativity. While much current research stresses innovation, there is also considerable interest in the act of creation itself, the processes underlying it, the characteristics of creative people and in how to elicit creative activity, more so because it has been suggested that there may be a common basis to creativity in both science and the arts.

Irving Taylor (1959) has analysed over a hundred definitions of creativity and found evidence for five levels. He suggests that it varies in depth and scope rather than type, and that it is misleading to distinguish between scientific and artistic creativity since it involves an approach to problems more basic than the accident of professional training. The first level is the expressive creativity found in the spontaneous drawings and musical compositions of children. It involves 'independent expression where skills, originality and quality of the product are unimportant'. At the next level of productive creativity there is a tendency to control free play and improve technique.

At the third inventive level discovery is an important characteristic, and involves flexibility in perceiving 'new and unusual relationships between previously separated parts'. The fourth level of innovative creativity is reached by few people; it involves significantly modifying the basic principles underlying a whole field of art and science.

The highest form of creative power is emergentive creativity, where an entirely new principle or assumption emerges at a fundamental and abstract level.

The composing of music and poetry does indeed seem to be psychologically akin to solving a difficult mechanical or mathematical problem or to formulating a new hypothesis in science.

Four stages seem to be commonly experienced in problem solving:

1. The initial appreciation of the existence of the problem, the taking stock of the situation and the trying out of unsuccessful ideas.

2. The incubation period when the problem has apparently been put into the 'back of the mind', the person concerned being asleep or even engaged on some other activity.

3. The inspiration stage when the solution is suddenly seen.

4. The elaboration stage where the solution is criticised or experimentally tested or the details of the idea worked out. Such stages are often interwoven and may appear to merge in the continuing dynamic processes which occur during creation.

Helen Durkin carried out some interesting experiments at Columbia University before the war. She gave her subjects various problems to solve and asked them to think aloud while so doing. The period just before the solution was reached seemed to be marked by a short pause of quiet intentness which sometimes involved an appearance of great tension and at others seemed to be merely a cessation of all visual activity. The tension seemed to be of suspense rather than of effort. This pause ended in 'an explosively expressed elation or in relieved relaxation'. Even when her subjects were not at all sure of the details, they seemed to feel convinced at this stage that a solution had been reached.

Of course not all problem solving ends with an intense sudden 'Aha!' of inspiration. Insights can often be only partial. 'Inspired' solutions may turn out to be wrong. As W. Gillies Whittaker, well-known particularly for his choral works and setting of English folk songs, tells us: 'In the warmth of inspiration one often has the impression that the products are of outstanding merit and one is delighted with them, but on careful consideration after the enthusiasm has waned, they may appear valueless and need to be put on the fire at once' (see Whittaker, Hutchison and Pickford, 1942). John Eccles, the neurologist, in his article *The Physiology of Imagination* confesses that his one sudden illumination was proved several years later to be wrong.

To describe inspiration in terms of the Unconscious is, as Vernon (1931) remarked, 'no more enlightening than calling it "a gift from God" '. However, the term 'unconscious' need not imply Freudian connotations. We may ask with Eccles (1958) 'What activity in the brain corresponds to the creative activity of the subconscious mind, and how does it eventually flash into consciousness?' Eccles suggests that some failure in the synthesis of the engrams or some conflict in their inter-relationship is the neuronal counterpart of a problem that clamours for solution. The subconscious operation of the mind involves intense and unimaginably complex interplay of the engrams. New patterns can be expected to arise from the progressive change in the patterns of the engrams resulting particularly from interaction with other patterns. Should an emergent pattern combine and transcend the existent patterns it may, Eccles believes, produce some resonant-like intensification of activity in the cortex, which will bring this pattern to conscious attention, when it comes to light as a new idea.

Recognition of the part played by subconscious 'inspiration' in composing music does not mean that we underrate the importance of the assimilation of the music of other composers. As McGeoch (1942) says of problem solving:

> Where the subject 'sees into' the fundamental relations of a problem or has insight, transfer seems to be a major contributing condition. It is, likewise, a basic factor in originality, the original and creative person having, among other things, unusual sensitivity to the applicability of the already known to new problem situations.

This latter point, the originality of the composer, is one that Revesz (1953) stresses as the '*real problem* of creative activity'. However important recollections and analogies are in the creative process 'significant inspirations are independent of already familiar forms to a very great degree' (Revesz, 1953, p. 203).

In physiological terms not only must the creative brain accumulate an immense wealth of engrams of a highly specific character (i.e. the creative musician must have absorbed the music of many other composers), but it must also possess, in Eccles's words, 'a particular potency for unresting activity' in order to produce original works.

Emotion is commonly thought to play an essential part in musical composition. Indeed artistic creation may possibly be distinguished from scientific insight by the greater amount of emotion it involves. Bahle (1934), for example, gave 30 composers a number of poetic texts, asking each to select one, to set it to music and then report his introspections. He concluded that the poem must evoke an emotional response which is followed by an attempt to express this in music. Music, therefore, appeared to be an expression of emotion, even under the somewhat artificial conditions of the experiment. It is usual to add that the 'emotion must be "recollected in tranquility" or psychologically "distanced" or in some way digested or assimilated' (Howes, 1958). Tchaikovsky's 'Pathetic' Symphony is sometimes felt to be too 'close' an expression of emotion, i.e. that its emotion is more 'emozionalità non elaborata esteticamente' (to use Croce's words). The chanting of children (see Chapter VII) or of savages would seem unsatisfactory aesthetically for a similar reason. In Wagner's words 'When a musician feels prompted to sketch the smallest composition, he owes it simply to the stimulus of a feeling that usurps his whole being at the hour of conception.' The passionate and lasting emotions which drive the musician to write an Eroica Symphony 'may date from outer causes . . . but *when* they force the musician to produce, these greater moods have already turned to music in him' (Wagner, 1841, cited Howes, 1958). It may also sometimes happen that ideas that have been 'simmering' in the composer's brain are given overt expression as soon as some external stimulation occurs – be it some emotional event or the commissioning of a work by the BBC.

Hebb's hypothesis on emotion might seem to provide a clue to how emotion might possibly function in giving rise to new compositions: 'Strong emotional disturbance tends to prevent the repetition of any line of thought that leads up to it' (Hebb, 1949). If some such account of emotion is true, perhaps it might act by clearing the composer's mind of old solutions to musical problems, thus enabling new combinations necessary for original compositions to occur. But the emotion need not be a 'disruptive' one. One night on a voyage to Australia, Whittaker was deeply moved by the beauty of the sea and sky. The next morning he came across William Habington's poem *The Coelestriall Sphaere*, read it three or four times, went to his cabin for paper and started writing immediately. Whittaker considered his setting of the poem for chorus and orchestra to be one of his most successful compositions.

After the 'storm' or crisis of the emotion has passed, the ideas may be produced so freely that, as Elgar says, composition seems merely a matter of taking as much of the 'music in the air' as the composer requires. During the actual process of composition, emotion may enter again, now apparently more as a concomitant than as a causative agent, perhaps as a feeling of joy at having found an expression for the previous emotion or in working out ideas and constructing a movement. The composer is often described as being in some kind of state of clairvoyance. What Mozart calls a 'pleasing lively dream' may resemble hypnosis, rather than sleep. According to electro-encephalographic evidence (Grey Walter, 1953) in hypnosis 'awareness is not lost, but heightened – restricted, it is true, to specific categories of stimuli, usually the hypnotist's voice' (Beethoven's 'inner voice'?). The basic properties of brain function remain undistorted, 'so that all the pathways of association and stores of experience are intact'. The difficulty of checking such a supposition experimentally would be, of course, finding a composer whose inspiration did not desert him in laboratory conditions. The most creative of Walter's colleagues had brain-waves characterised by a rich variety of patterns. One might at least check this point among creative musicians.

Creative ability is very hard to test, even by individual methods (Wing, 1948). Vater (1934) and Vidor (1931) gave their subjects a tapped time pattern on which to build a tune. Such tunes could be sorted out, Wing suggests, into typical shapes, on similar lines to the drawings of a man at various ages.

Some of the methods used to test 'creativity' in general abilities might be adapted to music. In recent years, psychologists such as Torrance and Taylor in America, have given considerable thought to means of identifying individuals who possess high ability for original thought. The traditional type of intelligence test where there is only one right answer seems to do less than justice to such individuals. Creativity tests are designed to encourage inventiveness and originality. Typical questions are: 'How many uses can you think of for a brick?', 'What would be the consequences if the world became 10 degrees warmer?' Children who do well on intelligence tests are thought to be more willing to conform to ideas of their classmates and teachers, in contrast to 'creative' children whose ideas tend to 'diverge'. However, Liam Hudson (1966) has used both types of test on boys at a school where they were encouraged to produce original solutions to technological problems. Some of the most inventive boys did not score particularly well on either test. Hudson believed that this may have been because the tests did not really evoke their interest. He refers to results of American research which suggests that the ability to channel one's interests even to an obsessive degree may be a condition for producing original work. That is certainly likely to be true of composers of music. It is also probable that difficulties might arise in interesting potentially highly creative musicians in 'creative' tests. Revesz's approach of attempting to study spontaneous compositions of individuals with marked creative talent from the point of view of originality and development of individual style might be more promising than setting tests.

The composer certainly requires a high degree of musical ability to build up the rich deposits of engrams as raw material with which his own mind can work. According to the American composer, Henry Cowell (1926), in order to compose seriously the composer must have a type of mind that is capable of thinking as correctly in terms of sound as a literary author might think in terms of words. Since he was unable to attend as many concerts as he would have liked as a child he formed the habit of deliberately rehearsing afterwards the compositions that pleased him. He soon began to experience 'glorious sounds leaping unexpectedly into his mind'. With great effort he succeeded in keeping some kind of control over these musical ideas until he became able 'to produce more and more readily whatever melodies and harmonies and tone-qualities' he desired.

Strong auditory imagery seems to be required for effective composition. Agnew (1922) examined accounts of the auditory imagery of Schumann, Mozart, Berlioz, Tchaikovsky and Wagner. Schumann possessed powers of vivid, accurate and profuse tonal imagery. Berlioz could hear his compositions mentally. He even dreamed of themes and wrote them down when he awoke. Musical sounds seemed to follow Tchaikovsky wherever he went and whatever he was doing. Melodies came to him fully harmonised.

EXECUTIVE TALENT

Many composers have shown, as did, for example, Chopin and Liszt, considerable talent as instrumentalists. How far this is a matter of two distinct abilities being combined in one individual or of one ability being shown in more than one direction is difficult to judge. Revesz believed that creative talent and interpretative talent were distinct and, rightly, points out that virtuoso performers rarely produce original compositions of much worth. This might be largely due to the demands made on their time and energy by the exacting nature of their profession.

Personality factors may also partly determine whether an individual turns to performance or composition. Wing (1941*a*) believed composers tend to be introverts, and performers, extroverts. It is noteworthy that extroverts (classified as such on the basis of a personality inventory) have been found to have quicker reaction times than introverts (see Broadbent, 1958).

It seems likely that creative talent would be closely connected with the interpretative side of performance, if not with the technical side. Like the composer, the interpretative artist seems to draw upon subconscious sources for the solution of his problems. Thus Bruno Walter (1961) refers to the 'psychological phenomenon that must be familiar to every gifted re-creative musician. Whenever I had been in doubt for some time about the right speed for a musical phrase or episode, it happened that I was suddenly faced with a decision coming, as it were, from a deep region of my mind; as in a moment of revelation the right speed had dawned on me, giving me a feeling of complete certainty.'

As we noted in Chapter II, tests of aptitude for the executive side of performance have received much less attention than those for the aural side. Seashore (1919) did include tests of strength of grip, precision of movement and timed action, among those he thought necessary for a complete talent chart of the prospective music student. But like his measures of sensory capacities they were probably on too elementary a level to be useful. Whittington (1957) found only very low correlation between musical age (as measured by the Wing tests) and manual dexterity tests. Kwalwasser (1955), however, cited a number of studies where sizeable correlations had been found between his tests and various measures of motor abilities and muscular control. How far this was due to those who scored high on both tests having better muscular co-ordination due to training on a musical instrument is not clear. Whittington's subject who played the piano often fumbled when fitting blocks together in the manual dexterity task. But this whole question requires further research. Wing (1948) was probably right to call the aural ability required for his test a pre-requisite, but no guarantee, of performing ability. In any case, as suggested in Chapters VII and XX, the auditory and kinaesthetic sides of musical ability appear to be intimately connected from a very early age. Provisional figures suggest that Thackray's performance tests are quite closely related to his perceptual tests.

'MUSICALITY' (AESTHETIC APPRECIATION OF MUSIC)

'Musicality, primarily, denotes the ability to enjoy music aesthetically' (Revesz, 1953). In his view, musicality is to be distinguished from (*a*) creative or interpretative talent, (*b*) aural abilities (such as capacity to discriminate rhythmical and tonal relationships) and (*c*) affective response to music, or love of, or interest in, music. While recognising that there are various degrees of musicality from the very pronouncedly musical downwards, Revesz states that 'to be considered musical a person must possess several' of these characteristics:

> Ability to contemplate a piece of music as a work of art, to assimilate it and to co-ordinate what he has heard, following the parts without separating them from the whole; to be sensitive to the artistic quality of composition and performance; to understand the structure of the work; to follow, even to anticipate, the composer's intentions; and to become

so absorbed in the emotions expressed that he feels as if he were creating it.

For Hevner and Mueller, too, the aesthetic response to music is highly attentional, with every detail being followed and 'making the experience of it a forceful and vivid awareness' (Hevner, 1937). It involves a keen perception of the music's qualities, of rhythm, harmony, melody or all these elements combined, without which its beauties may be completely lost. The affective accompaniment gives the response importance and significance. The feeling reaction is not so intense that it completely absorbs the listener's attention, but may be sustained over a long period. According to Hevner and Mueller, it is the background of widespread and unlocalised bodily sensations, especially from the involuntary muscles and viscera which give the experience affective and emotional qualities.

Vernon (1931) emphasises that there is

> no one standard experience which can be called the aesthetic but that it is a synthesis of all the various tendencies, different for every individual. It should include as many diverse elements as possible and it is the presence of overmuch attention to any one aspect (be it intellectual, emotional, gregarious, or anything else) to the detriment of the whole that really constitutes the non-aesthetic factor in musical appreciation.

In the highest moments 'which occur but seldom' the various elements 'integrate into a total cognitive-affective experience'. Support for this view was shown by the rankings for aesthetic appeal of items played at two experimental concerts. These did not run parallel either with the marks for intellectual interest, nor with those for emotion, but were rather closely parallel to the sum of both.

Though Revesz distinguishes between musicality and musical ability, the appreciation of form would seem to imply ability to recognise a theme when transposed or rhythmically or melodically changed (i.e. the ability tested by Drake's Memory test). Sensitivity to the finer points of style would seem to require the sort of ability tested by the Oregon or by Wing's last four tests. The last tend to be distinguished from the aural acuity tests of Wing's battery in only a relatively minor factor (see p. 207). Crickmore (private communication) seems to be the only person who has tried to obtain a quantitative measure of aesthetic listening and to compare it with musical ability. In his experiments he assumed that if a musical composition has been assimilated aesthetically it will leave the listener feeling interested; happier; more relaxed; and with a desire to remain quiet; satisfied; and without any particular mental pictures. He asked two groups of engineering students attending his music appreciation class as a Liberal Study to rate themselves on each of these points immediately after listening to records (cf. p. 192). The number of syndromes (complete sets of 'correct' scores) for each individual was compared with his Wing score and with his (self-rated) liking for music. In each case the correlations obtained were low. It is of course highly probable that any testing situation evokes an analytical, rather than an aesthetic, attitude on the part of the listener.

It would certainly seem wrong to deny the genuineness of aesthetic experience of persons with only a modest amount of musical ability. Grillparzer's character of the violinist in *Der Arme Spielmann* who delighted in sustaining a single note, then alternating it with the fourth, the fifth and the third above (Bruno Walter, 1961) was possibly as highly attentive, as keenly perceptive and as emotionally responsive to this very simple musical activity as a trained musician to a Bach fugue.

Myers (1922) believed that the experience of beauty was probably rooted in man's remote past when it could be evoked by such simple material as one or two tones or splashes of colour, by the most primitive *forms* conceivable of art material, just as today it is evoked by more complex forms.

A high degree of musical ability is conversely no guarantee of a high degree of aesthetic appreciation on any particular occasion. Aesthetic appreciation has something of the elusive quality of the composer's 'inspiration'. A concert-goer cannot say 'When I hear Beethoven's *Missa Solemnis* performed tonight I will have a profound aesthetic experience' any more than a composer can say 'I will have an inspiration for a new symphony after breakfast tomorrow morning'. Just as the composer may owe his inspiration for a composition to the effect of some strong emotion, so some emotional crisis in the life of the listener may result in a deeper degree of aesthetic appreciation. At a time of great sorrow, the mundane cares of our everyday pursuits may pale into insignificance, leaving our eyes and ears free to behold the beauties of nature and art.

Various attempts have been made to study the emotions experienced while listening to music. For example, verbal reports by the listeners have been compared with physiological records of heart beat, of pulse rate and of the galvanic skin response. (This last is a change in the electrical resistance of the skin which may be associated with an emotional state.) Usually some positive indications of emotional responses, especially among subjects who are interested in music, are obtained. (See Lundin, 1967, and Farnsworth, 1958, for a more detailed discussion.) Whether many listeners experienced any deep aesthetic feelings under the conditions of such experiments is open to question.

The possibility of inducing aesthetic experiences by taking such drugs as mescalin and LSD has received considerable publicity of late. In an experiment with mescalin Aldous Huxley found that his response to music was in no way comparable with the remarkable visual perceptions which he experienced. But, he asked, 'Would a naturally gifted musician hear the revelations which, for me, have been exclusively visual?' Ronald Cleak, an able and experienced musician, determined to try. Having taken a quarter-grain of mescalin, he sat down to listen to music in a room with a view of the sea. He found that his perception of colours was raised to a higher power and that subtle differences of shade and colour, which would normally pass unnoticed, became very clear. His perception of music, however, became grossly distorted. The first movement from the '*New World*' Symphony sounded hopelessly over-dramatic and at times even crude. The main subject in Bach's *Fugue in C sharp minor* from Book I of the 48 stood out unduly, the effect of the modulations was often lost, resulting in an over-romanticised and distorted version. In Schumann's *D minor symphony* perception was so distorted that listening became unbearable. Cleak concluded the emotional response resulting from drugs was no more conducive to a balanced aesthetic response than is a response in a person who has little natural emotional response to music. Composers of the stature of Beethoven, Mozart and Brahms might possess a superior ability to react emotionally to musical patterns, compared with the average musician. Such a capacity might have a biochemical basis, but must be of a much more subtle nature than can be induced by the taking of a drug.

In so far as the listener may enter into the music so completely that he feels he is creating it, it might seem that the creativity of the composer differs in degree rather than in kind from that of the 'musical' listener. At least some of the distinction between musicality and creative or interpretative talent might be attributable to lack of opportunity to learn to compose. There is certainly an enormous difference in opportunity for the child to learn to express himself in music as opposed to words. However, in spite of all the attention given to composition in the vernacular, the number of persons becoming great writers or poets remains small. The great variety of literary productions may fit into a continuum from the lowest to highest, with, as Burt (1943) suggests, a few individuals at the top showing really outstanding achievement. The comparative lack of opportunity to compose music may result in fewer compositions of lesser merit being produced, rather than affecting the higher levels of creative achievement.

XXIII

Ability in Music and Other Abilities

How far is musical ability related to general intellectual ability and to academic attainment? The 'A' stream of a school is usually better at music than the 'C' stream. Yet many highly intelligent people do not seem to be able to hear the difference between *God Save the Queen* and *The Star Spangled Banner*, while some mental defectives can play by ear. This question has important educational implications, which will be discussed in Chapter XXV.

Other questions to be considered in the present chapter are: Is there a broad ability which embraces all the arts? Is there any basis for the persistent notion that musical and mathematical abilities go together?

MUSICAL ABILITY AND INTELLIGENCE

The results various investigators obtained by comparing intelligence tests scores with scores of musical ability tests have been tabulated in Appendix III.

Most of the correlations are positive, but low. Both Kwalwasser and Wing refer to ·30 as being about the correlation to be expected with ordinary unselected objects.

When calculating his coefficients, Wing (1948) observed that there was good agreement between low intelligence and low scores on his tests, but that disagreement occurred where a high IQ was accompanied by a low musical ability score. Edmunds (1960), also, found that low intelligence and low musical ability appear to be closely related. When a certain level of general ability is reached, approximately IQ 90 for children of secondary school age, intelligence no longer plays a significant part, i.e. children may be musical or unmusical. Understandably, Coulthard's investigation with grammar school children produced zero correlations, and many of the studies based on testing College Students result in low correlations. With the RMSM boys in Shuter's results, correlations higher than ·2 might have been expected, since some 80% of the boys came from secondary modern schools. Two selection effects may have been involved. Firstly, only successful recruits were included and secondly, the least musical of the group were those most likely to owe their acceptance for training to other compensatory qualities of general ability or character. Edmunds believed that his results verified Burt's (1958) observation that the ability of backward children at music appears unusually high only when compared with their performance in more academic subjects and that rarely are they as good as the normal child of the same age. However, some intellectually dull children do seem to have average or superior musical aptitude. As Wing (1955) points out, it is important that such children receive suitable opportunities for developing their musical gifts. Striking, and if we are honest, puzzling instances of unexpectedly high musical aptitude occurring in individuals of extremely low general ability are the idiot savants. As we noted in Chapter IX, many idiot savants exhibit an outstanding memory for tunes, no doubt partly due to a concentration of interest, sometimes obsessionally, on music. Many also show superior powers of recalling events and dates.

At the lower levels of intelligence, the connection between intelligence and musical ability is probably at least partly due to adaptation to the testing situation, willingness or capacity to concentrate and so forth, but is there a connection at a higher level? Wing's view of the matter is that 'musical intuition' (i.e. the rapid mental understanding of the music or musical tasks) may be regarded as a form of intelligence 'although it may not be adequately measured in the normal intelligence tests which deal with logical reasoning'. McLeish (1950) found that 'speed at the higher levels measured by Cattell's timed intelligence test has an appreciable influence on the Seashore memory and pitch tests'. Speed might be expected to be even more important to music than, for example, to mental arithmetic; for while many other tasks can be practised at the learner's own speed or with erratic variations, too slow or irregular a tempo quickly destroys the character of the music. Musical ability must certainly involve what Hearnshaw calls 'temporal integration'. Hearnshaw (1951) proposed that 'perhaps the most basic intellectual skill is that of transcending time, involving the temporal integration of the stream of experience'. Hearnshaw later described several experimental tests involving generalisation over a series of successively presented data. Whether these would show a higher correlation with musical aptitude than present intelligence tests would have to be a matter for further research. As we noted in Chapter IV, some connection may exist between a deficiency in dealing with certain tests of time and retarded development of speech. There seems to be but little connection between Seashore's time test and intelligence, but this would not preclude there being a minimum level of 'temporal integration' required for normal speech development.

Franklin believed that in his second study he had found 'an indication that the higher the nature of the musical functions, the further we passed away from the sensory level and the more we approached a level where general intelligence and special intelligence join forces mutually supporting each other. In the case of the great creative artists, the great composers, conductors, musical theoreticians and writers on music this must undoubtedly be the case.' Indeed, have we not all at times wondered with Valentine (1962) whether the ability to follow and enjoy a Bach Fugue is quite unconnected with general intellectual ability? Franklin's own evidence is in fact quite weak, i.e. the strong loading on his TMT group test in a factor slightly loaded on intelligence. Franklin claimed that his TMT group tests required 'an evaluating judgement with regard to a synthesised collection of sensory impressions of different kinds'.

He argues further that 'investigations concerning the connection between musical and general intelligence must be made at the same level' and not, for example, by comparing exclusively sensory tests with intelligence tests. Verbal talent, he suggested, should be correlated with phrasing ability or something similar to it.

That the musically great men of history did possess far better than average intelligence is well established. In her meticulous study of the biographies of great men, Cox (1926) included 11 musicians. Their intelligence was estimated from evidence of the activities of which they were capable at various ages. Bach's IQ was thought to lie between 125 and 140; Beethoven's between 135 and 140, and Mozart's between 150 and 155. Taken as a group, however, the musicians were among the lowest in IQ of all the eminent men whose biographies she studied. A later investigation of their versatility also showed that they were among the least versatile (White, 1931).

An interesting attempt was made by James Roderick (1965) of Illinois University to find out whether the Minnesota Tests of Creative Thinking would show any particular connection with musical ability. He tested two groups whose main subject was music, one whose speciality was art, and one randomly selected college group. The relationship between the Wing and the Aliferis tests, and the creative ability tests was no greater than on intelligence (range of correlation ·20 to ·35). There was no connection at all between the Drake Memory test and creative ability and no significant difference between the music students and those studying other subjects. After attending a musical course, both music groups improved in *fluency* of response to a hypothetical music education problem, but not in *flexibility* of response.

Considering the abstract nature of music we might expect non-verbal intelligence tests to be more closely related to music tests than are verbal ones. However, only to a minor and inconsistent extent is any such tendency apparent in Lundin's or in Gordon's results.

We may also ask 'Are some aspects of musical ability more closely related to general intelligence than others?' Again our table does not show many consistent trends. This may be partly because many of the subtests are much less reliable when taken separately. Memory tests sometimes give, as might be expected, relatively high correlations, but pitch tests are often slightly higher still. In an early research, Burt (1909) found sizeable correlations between intelligence and pitch discrimination, as tested with tuning forks. He thought a possible explanation might be that the development of intelligence in man depended upon the power of speech which in turn partly depended on auditory discrimination.

Some tests of rhythmic aspects of music, for example, Holmstrom's and Gordon's, show a moderate connection with intelligence. This agrees with the opinion expressed by Botez and Wertheim (1959) that rhythm is a function of superior integration the limits of which are much wider than musical functions (cf. p. 191).

MUSICAL ABILITY AND EDUCATIONAL ACHIEVEMENT

In 1937 Mursell concluded from the data then available that when functional criteria of musicality are involved, such as teachers' estimates or marks for school music, musical ability may show quite a close association with educational attainment. He quotes examples of Continental investigations which did seem to suggest that all-round ability usually included musical talent. Some similar evidence was quoted by Doron Antrim (1945). During a period of 30 years at Magdalen College, Oxford, while only 10% of the students studied music, the music students won 75% of all the prizes and scholarships offered by the College. Antrim also stated that the IQs for music students at the New York City High School of Music and Arts averaged 11% higher than the general level for students in other New York High Schools.

It is difficult, however, to assess how far personality factors like willingness to work hard influence such results. Moreover, teachers' marks are liable to be highly saturated with what is known as the 'halo effect' of the pupils' other school work. The halo effect is a term used for the tendency to generalise from achievement in one sphere.

If we judge by the type of secondary school attended (cf. Chapter XV), grammar school children tend to make higher scores on the Wing and Bentley tests. This may be partly due to greater powers of concentration on such a task as taking an examination or a musical ability test. Again, as Cleak (1958) pointed out, a grammar school class are likely to achieve much better results at singing than secondary modern children, partly because of a fuller appreciation of the meaning of the words and of the need to sing expressively, as well as a surer knowledge of musical notation.

When Drake (1940) studied the relationship between scores on his memory test and the college grades of women students, he found correlations ranging from —·13 for Social Science to ·24 for Chemistry. In another study a correlation between his music memory test and a score representing the total credits earned by each of 230 students was only ·16. A zero correlation was found between his rhythm test and cumulative grade points (Drake, 1957). Wing (1954), too, reported low correlations of a similar order to those with intelligence tests between School Certificate marks and a number of subjects and his own tests.

Holmstrom (see Appendix IV) found that the factor associated with success in the three R's was almost wholly independent of the musical ability tests. As might be expected, *knowledge* of music was related to it in all groups, except for those children who came from homes where there was little interest in music or music-making.

However, Holmes (1954*a*) found that 19% of the variations from the mean on spelling test scores seemed to be associated with some of the musical abilities tested by the K–D tests. Pelletier (1963) investigated whether instruction on a musical instrument would facilitate the development of reading abilities among nine-year-old children. He constructed a preparatory fiddle which, like Fred's (see p. 249), could be taught by a teacher without experience of stringed instruments. After about six months of instruction on this instrument, the reading comprehension, but not the reading vocabulary nor the spelling, of his experimental group had improved significantly over a control group. The poorest readers had benefited most. Reading and spelling were related to tests of pitch and of time, but not to melodic memory.

MUSICAL ABILITY AND OTHER ABILITIES

The table in Appendix IV shows correlation coefficients obtained by comparing tests of musical ability with tests of various other abilities.

Only oral French shows any marked correlation with musical ability. Although Coulthard's subjects numbered only 32, his test of oral French was quite an extensive one, including subtests of pronunciation, accent, intonation, phrasing and fluency; it lasted altogether over two and a half hours. His results confirm the popular view that musical children have an advantage when it comes to learning to speak a foreign language.

In his larger study, Franklin found that a factor which had strong loadings in the tests of visual form perception also had small, hardly significant, loadings in the two tests of melodic rhythm (Wing's and Franklin's own). Franklin suggests that since rhythm is very 'important for musical form, these small loadings might indicate something in common for the perception of visual and musical form'. The results of Karlin's study, which included visual as well as auditory tests (see Appendix II), suggested that certain speed and memory span factors may be common to both modalities. The aural form of Thackray's rhythmic tests correlated ·46 with the visual form (total scores). Two subtests showed moderate correlations between the equivalent forms in the two modalities: Accents (·45) and Rhythmic Patterns (·51) (Thackray, 1966).

Musical ability might seem to be especially connected with ability in the other arts; traditionally a connection with mathematical aptitude is said to exist. Findings in these two areas are discussed below.

MUSICAL ABILITY AND OTHER ARTS

One of the conclusions which Feis (1910) drew from his data on the genealogy of great musicians is that their parents have often been distinguished in the other arts or literature. This suggested to Mursell and Glenn (1931) that distinctive musical talent is a manifestation of a high level of all-round ability and particularly of a high level of artistic and literary ability.

The actual correlations of tests are, however, low. Both Carroll (1932) and Rigg (1937) used the Oregon music discrimination test with which to compare literary tests of their own devising. In Carroll's prose test each item contained four short passages which the subjects had to rank in order of merit. The passages were drawn from four sharply differentiated sources: classical literature, poor quality prose, pulp magazines and a deliberately distorted version of a passage of similar content. Rigg's subjects had to compare each of 35 short selections from works of standard poets with an inferior parody. Williams, Winter and Wood (1938) conducted their investigation under the supervision of Burt. They used three group tests of musical appreciation. In the first, the children were given a list of ten familiar tunes and asked to rank them in order of choice. Secondly, ten pairs of gramophone records were played, the tunes in each pair being similar in character but of very different merit. The children were asked which of the two extracts they liked better. In the third test, three versions of the same brief piece were played on the piano, the children being asked which was the best and which was the worst. With general intelligence held constant, the correlation remaining between the literary tests and the music tests was barely significant. However, Williams, Winter and Wood concluded that their results largely confirmed

> Burt's inference from his earlier test results that in criticism, if not in creation, in the vast majority of persons, if not in the specialist, and in the young if not in the old, aesthetic appreciation is dependent upon a group factor common to all the various media as well as upon general intelligence and special capacities peculiar to the different forms.

Burt himself claimed that he had obtained evidence of a 'moderately large' general factor for the appreciation of music, painting and literature. Unfortunately details have not been published (see Valentine, 1962, for a brief account).

Burt suggested that the common factor between the three arts might depend on the appreciation of significant form involving the apprehension of the relations between various elements. In the writer's opinion a musical composition may certainly evoke a similar aesthetic experience as a poem or as a painting. On the cognitive side of musical abilities the writer is inclined to agree with Guilford, the American psychologist, who has carried out extensive investigations of the higher intellectual abilities. In an article published in 1957 on artistic abilities, Guilford pointed out that evidence existed that auditory memory was distinct from visual memory. He postulated that similar distinctions might be found between the abilities to produce and express ideas in the graphic arts and the parallel abilities in music. Though he considers these may be distinct, they may not necessarily be independent or uncorrelated. In fact, he suspects that there *is* something in common among the parallel factors in the different arts (apart from the coincidence of some individuals having talents for more than one). He proposes, however, first to measure the factors for the different media separately, then to investigate any inter-correlations.

It is possible that there may be a correlation among artists of different types arising from personality rather than cognitive factors. Charlotte Bühler (1935) concluded from her extensive studies in Vienna of young children that they will take up any art that happens to be readily available in the environment as a means of expressing themselves. It does not follow, however, that individuals would be equally good at any art.

MUSICAL ABILITY AND MATHEMATICAL/SCIENTIFIC ABILITIES

In the words of Frank Howes (1958)

> The analogy between mathematics and music has been recognised from antiquity, and though all attempts to press the analogy, or even to define it, soon break down, it is still recognised by musicians and mathematicians and the rest of us who are neither . . . as a way of thinking in relationships, abstractions – there is an obvious similarity . . .

Thinking-cum-feeling in formal patterns of measurable units and relations, he goes on to say, appears to be a good description of a sonata movement as of mathematicizing.

One difficulty which arises in trying to investigate the connection between the two abilities by testing is due to the very different treatment mathematics receives as a school subject. Considerable attention is given to arithmetic and other branches of mathematics and its utilitarian value as an examination subject is made apparent. Music, except for the very talented, may seem to have much less importance, except as a hobby. For the very talented, music requires a considerable amount of time – sometimes at the expense of other studies. Even if they had much aptitude for mathematics, many musicians may have had little time or opportunity to develop it. Revesz (1953) found only 9% of professional musicians had mathematical talent or interest in mathematics. Some of the apparent lack of aptitude may have been due to lack of opportunity, or even to the attitude of being 'above' practical, everyday affairs affected or genuinely felt by some musicians, and the lack of interest due to mathematics not entering very largely into the popular hobbies. One wonders what percentage of the population of comparable socio-economic status to the musicians *would* express an interest in mathematics.

The results of correlating music and mathematics tests certainly do not reveal any close connection apart from the correlation of ·41 between the Seashore Memory test and a Number Series test.[1] Wing (1954, p. 167) correlated performance on his tests with School Certificate results and came to the conclusion there appeared to be little relationship that could not be attributed to another common factor, e.g. memory, attention or general ability. Of all the correlations between the Wing tests and the Admiralty entrance tests calculated by Shuter, those for the mathematical tests were the lowest. Success in the latter could, however, have been partly due to good teaching rather than mathematical aptitude. Though such tests do, of course, measure number ability, computational efficiency is not always synonymous with a true understanding of number (see, for example, Williams, 1958). College grades in Chemistry and Mathematics showed higher correlations with Drake's Memory test (·24 and ·22 respectively) than did any other of 16 subjects, but the coefficients did not attain a statistically significant level (see also p. 231 above).

In order to find out if mathematicians were more musically gifted than members of other professions, Revesz (1946, 1953) sent a detailed questionnaire to over 500 Dutch mathematicians, physicists, physicians and writers. Among the questions on playing, singing, composing and concert-going, there were only six concerned with aural ability, and two of these related to absolute pitch. It is not clear how much weight Revesz gave the aural questions in assessing his results, which are shown in the table below:

	n.	Musical	Unmusical
Mathematicians	135	56%	44%
Physicists	172	67%	33%
Physicians	165	59%	41%
Writers	110	71%	29%

We should not perhaps consider Revesz's results entirely conclusive; it is questionable how far trying to assess musical ability by questionnaire is a valid procedure. The writers might have been inclined to answer less objectively than the scientists, though they might in fact

[1] A Number Series test involves seeing relationships between a series of numbers, e.g. 'Complete the following: 1 2 4 7 11 . . .' Perhaps this kind of number ability has some connection with musical aptitude.

be expected to be sensitive to nuances of sound. While the mathematicians no doubt all possessed a high degree of mathematical talent, the same is likely to be true of most of the physicists and some of the physicians.

Vernon quotes the following evidence of a positive relationship between music and mathematics: 'Of the 200 odd members of the Oxford University Music Club and Union during the year 1927–8, more than 60% were scientists (including mathematicians and medical students), while in the University as a whole the proportion was scarcely 15%' (Vernon, 1931, p. 117).

At first sight this might appear to mean no more than that scientists feel the need for a spare-time interest among the fine arts and chose music as offering the most contrast to their normal occupation. However, there might, in fact, be some real correlation of abilities to deal with abstract configurations which are involved in both mathematics and music. This might be particularly true of individuals who are highly integrated intellectually. A senior University might be expected to attract scientists of higher intellectual calibre and of more highly cultured backgrounds than a technical college. It would be interesting to know if technologists of lesser institutes also show a marked tendency to choose music as a spare-time pursuit.

CONCLUSIONS

The evidence so far available suggests that musical ability is largely specific. There is a fairly well established connection between general intelligence and musical ability in the case of younger and less intelligent children. It seems reasonable to interpret such a correlation in terms of some common ability, such as powers of attending, concentrating or following instructions. For the more intelligent the musical ability of the child depends more on the special musical factors than it does on his intelligence. Positive correlations are nearly always found between measures of musical ability and other cognitive aptitudes. However, except in the case of an oral test of French, the coefficients are low. Further research is needed before any connection with aesthetic ability in the other arts or with mathematical/scientific ability could be accepted as established.

XXIV

Conclusions

In music the elements of melody, time, rhythm, tone colour, harmony, variations of intensity and phrasing are combined to make an integrated whole. (In some modern music the ingredients seem to be presented only partially mixed and the listener left to do the integrating for himself!) The ability to apprehend music would seem therefore likely to be largely unitary.

Admittedly, it is possible to find compositions that can be considered predominantly melodic or rhythmic. Thus Henkin (1955) studied the preferences of college students for ten pieces of recorded music. A melodic and a rhythmic factor emerged. He was unable to find any recorded composition written in a 'purely harmonic idiom'. Consequently – and not surprisingly – he was unable to isolate a harmonic factor.

Similarly, it is of course true that most individuals find some aspects of music easier than others. For example, a violinist may have good intonation but have difficulty in playing in time, or a pianist may play rhythmically and with good phrasing and dynamics but find his pitch discrimination insensitive when he attempts to deal with stringed instruments, for example in conducting an orchestra. Fry's tune deaf subjects were often able to recognise a tune by its rhythm. Bentley reported no significant inferiority among 'monotones' at performing his rhythmic memory test. Mursell (1937), too, notes rhythm as an example of a capacity which alone might not make a musician, in the case of a jazz drummer who might be rhythmically effective, but tonally inept.

Teplov (1966, pp. 378–99) believes that there are three basic musical aptitudes:

1. A sense of tonality which enables us to sense the tonal relationships of the notes of a melody and the emotions expressed by melodic movement. This is closely connected with pitch discrimination.

2. The ability which enables us to reproduce a tune by singing and to play by ear, and to develop an 'inner ear' for music.

3. A sense of rhythm by which we are able to feel the rhythmic movement of the music and to reproduce it.

He thus divides the receptive and the reproductive side of music in the case of melody, but not in the case of rhythm. He describes three children whose varied talents between the ages of six and nine demonstrate that these aptitudes can exist independently. Nelly B. loved to listen to music, could complete unfinished melodies on the tonic and sing back tunes. She was also an active composer of music and of words for songs. However, her sense of rhythm was relatively poor. Her inability to keep in time when singing in a choir spoiled the singing of the class. Inna G., on the other hand, had a good sense of rhythm, and could sing back tunes perfectly. She composed with great facility, but showed little response to the expressive power of music. Adik G. loved to listen to music, had an excellent sense of rhythm and could recognise many tunes. Yet, he had great difficulty in singing in tune nor could he manage to play a tune by ear on the piano.

Teplov, however, agrees that musical aptitudes cannot exist independently. The various aspects of music are so intimately connected that a reasonable minimum level of all-round efficiency is required both for listening and playing. As Teplov says, without some sense of tonality and affective response to music, Inna would not have developed her 'inner ear'. We might go farther and say that unless Nelly B. can overcome her difficulty in keeping in time, her progress with any kind of performance, much less ensemble playing, is going to be severely impeded. Mursell's jazz drummer hardly deserves to be called a musician. Needless to add that many drummers have a high degree of musical ability.

Arguments similar to those of Vernon in connection with general intelligence (see Appendix II) might also be put forward for preferring solutions of factor analyses that produce a general musical ability factor. A general factor is likely to be more stable than group factors. (However, Holmstrom managed to find evidence for factors similar to his Alpha and Beta factors in several other studies.) Many subtests of musical ability are not very reliable when considered separately. In selection situations, as with intelligence, the first point to establish is probably the general level of the individual's musical ability. More specific abilities might be examined later. We do not deny that it is often of value to the teacher to have a precise analysis of his pupil's deficiencies.

The close connection between all stages of musical perception and memory also seems to support the idea of a broad general factor of musical ability. It is, however, plausible that pitch discrimination might form a separate narrow factor. The acceptance of an important general factor does not preclude minor factors.

It may be objected that such and such a child who has a poor score on a musical aptitude test is 'musical'. Further enquiry often reveals that what is meant is that he enjoys music. That is why it seems necessary to draw a clear distinction between being musical in the sense of having musical ability and being able to enjoy music aesthetically. We are also inclined to distinguish interest of a rather intellectual nature in music both from aesthetic enjoyment of music and from musical ability. The connection of aesthetic appreciation with creative and interpretative talent has not been empirically demonstrated but we suspect that it may be close.

To sum up, we believe that even individuals without musical ability may enjoy a deep and genuine aesthetic experience when listening to music. Without even a moderate level of general musical ability, many people may have some proficiency with one aspect of music, such as a talent for rhythm. Such persons may or may not be interested in music. The person with musical ability may have little interest in music, but if he has, he is likely to become a discriminating listener. If in addition he is proficient on the muscular side (or possesses a good voice), he has the basic equipment for mastering the techniques of performance. But to interpret music he also needs to be able to attune himself to the emotions expressed by the music he is playing. The composer would seem to require a particularly high talent for holding auditory images in his mind, as well as being able to play some instrument, but his peculiar gift is the ability to produce original ideas. In order to read music and be able to write musical notation, both performer and composer need a minimum amount of intelligence. Many performers and composers are in fact highly intelligent, but above a certain level, cognitive ability does not seem to be especially related to musical ability. Most of the empirical evidence points to the conclusion that musical ability is a rather highly specialised talent, though it may of course co-exist with other cognitive abilities.

All the above points could profitably bear further investigation. As techniques for studying personality improve, it would be interesting to study the connection between musical ability and various personality factors.

PART V

Music and Education

XXV

Music and Education

From our discussions of tests in Part One, we concluded that it is possible to arrive at a useful assessment of a child's musical quotient by the age of 11, or even earlier. Experience of using the more satisfactory types of test in prognostic situations has confirmed that they are of considerable help in determining which children are most likely to profit from instrumental lessons. Since, however, the muscular side of performance has been explored much less thoroughly than the perceptual, many children with promising test scores may lack the muscular co-ordination necessary for progress with the more exacting instruments. Again, many children with talent may have little interest in learning a major instrument. Sometimes being told that he has achieved a high score on a music test may be an incentive to a child to take an interest in music (Wing, 1948). Even if he rejects the opportunity of lessons at one stage, he may later on, perhaps through an interest in popular music, come to value his own gifts.

A most important question is – what can an individual with a musical quotient of x points achieve? As Farnsworth (1958) suggested, there is need to 'study more intensively the minimum level necessary for later success in several kinds of music skills'. The results of such research should take into account the possibilities offered by new teaching techniques. The programmed instruction to be described in the next chapter may be helpful especially to individuals whose progress is much faster or slower than that of the average of a class. Again, the Suzuki approach to the teaching of stringed instruments to very young children may make us wonder how much more could be achieved by the wider use of such methods, if they could be adapted to the social and educational systems of Western countries. Similar strategies might also be tried in the teaching of other instruments.

THE GIFTED CHILD

How far any individual should be advised to spend time on music depends on his musical ability and interest in music compared with whatever other interests and aptitudes he may have. Two classes of pupil deserve special attention to their musical education: (*a*) children with a high level of talent to whom the study of music could bring great pleasure, even if they do not take it up as a career; and (*b*) those of low general ability who have at least some aptitude for music.

Although musical ability appears to be largely specific, music might nevertheless, particularly for children whose musical aptitude surpasses their verbal ability, become a path into the wider avenue of Western, or indeed of world, culture. There seems less need since the development of radio and TV for quite so much emphasis to be placed on the written word as in the past. A music-centred education could be as valuable as one centred on language and literature, or on the sciences. For instance, the Elizabethan Age could be approached through the music of Byrd and the madrigalists as well as through Shakespeare and the dramatists. 'Team spirit', in so far as it can be taught at all, could be instilled just as well in the orchestra as on the playing field. The discipline and accuracy and the quick co-ordination between brain and body required for music also suggest an analogy with sport. The habits of concentration, patience, hard work and determination required by music are equal to those for any other subject. Intellectually the demands made by music have been compared to those associated with learning a new language in an unfamiliar alphabet (Martin Cooper, 1965).

Strong claims for the educational value of music among very young children were put forward by Antrim (1945) writing about Alexander Blackman's Orchestra which was composed of New York children, aged between two and six. During a period of 12 years over 800 children played in the orchestra. When they entered school all of them were found to be ahead of their classmates of the same age, the great majority being double promoted. How much of this advancement can really be attributed to their music training is, however, difficult to judge. Although Blackman did not select the children the parents who wanted their children to take advantage of this kind of experience probably provided an environment that was intellectually stimulating in other ways. While it may be true that music is an especially valuable activity for children under six, we may wonder whether other types of training might not prove equally valuable.

Examples of school music as a socialising influence are mentioned by

A. W. Rowe (1959) in his book on the education of the average child. When Headmaster of a secondary modern school, he allowed the children to bring their favourite records to play during the lunch hour on condition that they would also listen to records chosen by members of the staff. As a result of helping to organise a gramophone club, of joining a school harmonica group and of building guitars and a double bass, two of Rowe's pupils became considerably more co-operative and interested in their school work.

Providing a suitable education for the child whose musical talents are conspicuously higher than his other capacities is a relatively simple one. Provided that he himself and his parents are willing, the problem is basically one of finance. Thanks to the pioneer efforts of Ruth Railton, Yehudi Menuhin and many other less well known people, an appreciation of the importance of training gifted youngsters seems to be gaining ground in Britain. An intrinsic problem, however, is the fact that the person trained as a musician is qualified only for a musical career, whereas the linguist or scientist has a wide choice of careers open to him. One remedy is to combine music with a good general education, as is the aim of the Central Tutorial School for Young Musicians.

The musician whether he is going to become a performer or a teacher or both should have as good an education as can be fitted in with his musical studies. Among recommended qualifications for the College music teacher, Lee Chrisman (1962) of the University of Southern California included: 'knowledge of the humanities and/or the interrelationship of the arts'. These should be a vital part of the teacher's background. But many musical children are also above average in other subjects and could win university places which would lead to careers more financially rewarding than music. Most of the best players trained by Ruth Railton do not become professional musicians. One may hope that with the increase in the hours of leisure promised during the second half of the twentieth century, music will become ever more highly valued and the status of musicians both as regards money and conditions of work will improve.

The musical child with other talents is likely to be attending a grammar school or to be found in the top streams of comprehensive schools. These are the children most under pressure from the examination system. The problems of grammar school music were discussed at the Colston Research Society Symposium on Music in Education (Grant, 1963). Some grammar schools which are conspicuous in their success with university entrance and in sport are also excellent in their achievements in music. However, in many others music has a low place.

Noel Long quoted figures based on a survey of English grammar schools six years previously (but thought to be representative of the position in 1962). In about a third of such schools music was either absent from the curriculum, or else had little more than a token existence; nearly a half had no orchestra of any sort and as many as a quarter had no real choral training. In such a situation, Long believed that the paramount aim should be to discover the talented and to provide the conditions in which they can develop their talents as fully as the other school pressures will allow. By grading the pupils in such music classes as can exist in grammar schools, standards can be considerably raised. This in turn tends to improve the prestige of music in the eyes of staff, pupils and parents.

The musically talented children with good general ability are particularly important as providing a source of potential music teachers of the future. Many of those who lack the talent, interest or training to become specialist music teachers at the secondary school level, may still be capable of doing good work in music as primary school teachers. The use of musical ability tests by Colleges of Education would help to identify students who should be encouraged to specialise in primary school music and those who should be discouraged, because of a lack of talent.

If the students have come from schools where music has been neglected, they may well not be able to read music. The programmed instruction courses on the rudiments of music to be described in the next chapter could be of great help in teaching them to read music efficiently in a short time without the constant attention of the college staff.

More difficult may be helping students who are not 'monotones' in the sense of not being able to sing in unison with others, but whose intonation may stray when singing without a piano. HMSO Report *Music in Schools* (1960) claims that

> experience shows that there are few teachers of young children who cannot learn to sing simply, naturally and rhythmically, the traditional songs that are the children's heritage and should form the basis of all their musical training at the primary stage. Even if the teacher is a pianist she should cultivate the habit of unaccompanied singing, which allows her to get closer, both physically and imaginatively, to the children.

In many cases, practice and confidence may be all that is needed to make the teacher a proficient singer. However, many student teachers reach college without being able to match a tone or hold the pitch accurately when singing a familiar tune with an accompaniment. In such cases, Brody (1953) believes that the student must begin at the stage of the young infant by making his own pattern of sounds as he moves his body and learning to differentiate the sounds he is making. Only then is he ready to be helped to match his own tone to that of others and to find the patterns that he can sing on the piano. His progress, Brody warns, will be slower, not faster, than that of the infant. Because of the importance of this problem, colleges of education would seem to be particularly suitable places for research into how far pitch and other musical deficiencies can be overcome by suitable training.

THE 'AVERAGE' CHILD

If special provision is to be made for the talented what about the rest? The 'average' children are surely important, if only because there are so many more of them. The battle-cry of the Music Supervisors' National Conference in America is 'music for every child, every child for music'. In their '*Child's Bill of Rights in Music*', they included the following clauses:

> 1. Every child has the right to full and free opportunity to explore and develop his capacity in the field of music in such ways as may bring him happiness and a sense of well-being; stimulate his imagination and stir his creative activities and make him so responsive that he will cherish and seek to renew the fine feelings induced by music;
>
> 3. Every child has the right to make music through being guided and instructed in singing, in playing an instrument, and, as far as his powers and interests permit, in composing.

These are indeed fine ideals. But what about the child for whom music, far from being a source of happiness, may be, as is physical training to others, a form of torture (see Grant, 1962, p. 20)? Obviously music must not be forced on children, either by their parents or by their teachers. But quite possibly if such a child had been exposed to music in happy circumstances when he was younger, he would have grown up enjoying it. What about the child whose powers to make music are but slight? In general the case for the schools providing remedial training for 'monotones' would seem to be less strong than provision for backward readers, or for children with speech defects.

This question largely depends on how much effort must be expended on the non-singer to enable him to sing correctly. Because of the lack of real knowledge about the effectiveness of the various remedies suggested for inability to carry a tune, an extensive investigation is being carried out at the Western Illinois University by Oren Gould (1965). The first part of the project is a survey intended to collect information about the incidence of non-singing in the schools and on remedial methods in current use and their effectiveness. This will be followed by visits to the classrooms during which two diagnostic tests will be given to non-singers, as selected by their teachers. One test will check hearing with an audiometer. Then a 'speech and song response' test will record on tape for later evaluation each subject's responses to items selected to diagnose non-singing difficulties in the five categories of low speaking voice, lack of tonal memory, speech and vocal difficulties, pitch difficulties and psychological inhibitions. Remedial procedures will then be developed and tried out in a pilot study. After that it is hoped to train twelve teachers, one music specialist and one classroom teacher from each of six schools, to use the remedial procedures in their schools in experiments planned to last 18 weeks. Eventually Gould hopes to produce a manual of recommended classroom procedures for general use.

It is possible, then, that research on the overcoming of pitch deficiencies, including the use of teaching machines, may result in the development of effective means of dealing with the 'droners' which can be a source of trouble in the primary school. Even where remedial teaching is not feasible, the individual might be guided into musical activities where his weaknesses will not be too handicapping. If he is pitch deficient, he can try to play a percussion instrument, or one of the Carl Orff melodic instruments, or the piano. In any case, he must not be given the impression that because he has difficulty in singing in tune he will never enjoy music.

How far should schools aim to provide every child who wishes to learn an instrument the opportunity of doing so? The high percentage who begin to learn only to give up before making much progress has long been a matter of concern. Martignetti (1965) tried to find out why elementary school children give up music lessons. He collected information from 56 music teachers in New Jersey. The percentage of children who had been playing the previous year and had given up averaged 34%. The rate was higher, almost 50%, among beginners. According to the music teachers, loss of interest, due to a lack of support from parents, accounted for more than one half of the wastage. Lack of ability was a secondary factor. Martignetti interviewed 35 of the children and parents. 70% of the children said they found difficulty with the instrument they had tried to learn. 33% did not like the instrument and many believed their parents did not like it either. The parents gave lack of time for practice as a chief reason for their children giving up music.

Hal Bergan (1957) pursued a similar enquiry in five high schools in Michigan State. He compared the replies of 200 students who had given up playing with 200 who had continued. 62% of the dropouts were among the weaker players in their groups; the highest rate occurred among string players. Bergan, like Martignetti, stressed the need for co-operation between teacher and parents and for re-evaluation of selection methods. Martignetti believed that more attention should be given to matching pupil and instrument at the outset.

Being allowed to experiment with several instruments before lessons begin may serve both to stimulate the child's interest and suggest a suitable choice. For example, Long (Grant, 1962) recruited several promising brass players by setting aside one of each instrument for the boys to try during the morning break. One ten-year-old became a skilful and dedicated horn player. Previously he had been a poor recorder player and an unenthusiastic pianist.

The recorder may often serve as a useful exploratory instrument. In America the use of 'pre-instruments' seems to be very popular. Bernhart Fred (1956) of the Northwestern University developed a three-string instrument which could be taught by a class teacher who had no experience of playing a stringed instrument. Twenty-eight ten-year-old children enjoyed playing the instrument although such experience did not improve their performance on the Kotick and Torgerson and the Knuth[1] achievement tests more than did the normal classroom singing given to their classmates. But such experience might have been valuable if they had gone on to learn the violin.

A more extensive study of the results achieved from specific types of musical experiences was carried out by Richard Colwell and Glenns Rundell (1965). They matched three classes of 13-year-old children on Wing scores, intelligence and scholastic attainment. The control group continued to receive class instruction in singing. One experimental group learned the ukelele for a term. The other was assigned to a piano class where two students could use the piano at a time while the rest of the class had paper keyboards. All three groups showed a marked increase on the Knuth test and on a test Colwell had modelled on the harmony part of the Aliferis test. The improvement may have been largely due to the enthusiasm of their teacher for the experiment. A year later the same tests were administered. In general, the scores showed a

[1] A test requiring comparison of melodies heard with those seen in notation.

loss of about one-half of the gain made during the experiment. The piano students had retained most, and the ukelele group the least, with those who continued regular class singing coming in between. All three groups were superior to other classes of similar age in the school. The experiment did not produce any lasting change in attitude to music.

C. B. Nelson (1956) found a time lag between the end of training and an improvement in achievement test scores. His experimental group of ten-year-old pupils showed no significant improvement after a term in which half of their music periods was spent on instrumental study. A year after, their knowledge of musical notation had improved over that of the control group who had had no instrumental instruction. The experimental group at no time showed a significant difference in the type of music they preferred.

It would be unfair to judge the benefits that accrue from long-term study of a musical instrument from the above short experiments.

Writing about the teaching of music and art at a secondary school level Ralph Smith (1965) of Illinois University insists that everyone has the potential to benefit from instruction in music and art. However, any civilising effects will not be noted until art and music are experienced in sufficient depth over a prolonged period – by which he means six years.

THE MUSIC EDUCATION OF THE BLIND AND OF THE HANDICAPPED

In the special case of the blind, even the most unpromising pupils should be given all possible help to develop their musical potentialities.

The blind are sometimes supposed to develop superior powers in their other senses to compensate for their lack of sight. It is probably nearer the truth to say that the sighted do not use their auditory and tactile senses to the full. Seashore and Ling (1918) concluded from testing 15 blind students and 15 sighted high school students that 'the blind and the seeing are, on the whole, equally sensitive to the direction and intensity of sound'. The same was true of pitch. Kwalwasser (1955) tested 100 blind children with the eight K–D tests which do not require the use of notation. Compared with the seeing, the blind were only average in pitch discrimination, intensity and tonal movement. Their scores for tonal memory, quality, time and rhythm discrimination were somewhat better than average. Sakurabayashi, Sato and Uehara (1956) in Japan administered the Seashore Measures to 282 non-music students, to 148 music students and to 150 blind non-music students and to 17 blind students of music. The music students scored better than the non-music students, but no clear difference between the sighted and the blind was found.

Drake (private communication) tested the entire population of a blind academy. The average score for Musical Memory was very superior to his norms for sighted subjects. It would seem understandable that the blind should be superior in musical memory, rather than pitch discrimination considering their long experience of dealing with (and remembering) meaningful stimuli perceived in succession. However Derek Pitman (1965) found that 76 blind children aged 8–11 made significantly higher average scores on the Wing Chord Analysis and Pitch tests (i.e. analysis of simultaneous sounds and pitch discrimination). Their average scores on the appreciation tests were inferior to those of a group of sighted children, and their performance of the memory test was only somewhat superior. Heim (1963) tested 155 blind Americans, of whom 115 were aged over 17, with the Wing tests. The results he obtained were quite similar to Wing's English norms, except that there were rather more at the higher and lower extremes of talent.

Even if *all* blind students should receive some music education, it is still important to assess their musical aptitude, since a suitable course of instruction can more easily be provided if their ability is known.

The Wing test can now be given to groups of blind persons. Wing himself devised a piece of apparatus with which a blind child could record his answers. Pitman adapted the Taylor Arithmetic Frame for the same purpose. The children were able to register any number from one to ten, as required in the memory test. He found that one person could cope with up to 15 blind children and that as many as 50 could be tested at any one time if supervisory staff were available in the ratio of one teacher to five children. Kenneth Heim was himself a blind musician with 20 years' experience of teaching in secondary residential schools for the blind in the USA. He found that older blind children could write their answers to the Wing tests in Braille. For a minority of blind children, music may be an especially important study, since they may later be able to earn their living as performers, teachers, or piano tuners. But Heim also stresses the great value of music in the social and emotional development of the blind. For example, he mentions a boy who was very withdrawn on entering a residential school but who became much happier when, as a result of singing lessons, he gave a creditable performance at the school concert. Juliette Alvin (1965) in her excellent book gives many moving examples of music bringing joy to handicapped persons and suggests how best to present music to suit the needs of the various types of disability. The criterion of achievement in performance and appreciation of music in such instances is in terms of its therapeutic, rather than its musical, value.

'AESTHETIC' EDUCATION

Perhaps the most fundamental aim of all music teaching is developing the love of the beautiful in music. It would seem a basic requirement for a music teacher to feel confidence in the power of music to attract and in the child's capacity to respond. Now the nature of art and of beauty have long been the subject of philosophical controversy. Some may believe that there is something eternal and absolute about beauty; others may point to the influence of a particular culture on what is accepted as beautiful at any given date. Which side the individual teacher favours is not important. What does matter is that he should respect the individual pupil's vision of what is beautiful and lead him from there to deeper and wider views. The stimulus which evokes the cry 'How beautiful that is,' may, as Hevner (1937) remarks, be a landscape, a cathedral, a sonnet, a forward pass in sport, a lullaby – something as worn and trite as *The Road to Mandalay* or something crude and tentative that never before has touched a responsive cord. No matter; so long as the object has moved the child to awareness of its beauty, it can become the starting point on a journey to a wider appreciation of music.

What the teacher, or the music critic, or the 'expert' considers to be worthy of aesthetic contemplation, must not of course be presented as a set of standards to which the child is expected to conform. Rather, the teacher is like a mountain guide, who has seen the beauty of the heights and who has come down to the pupil's level in order to lead him upwards. Many will not wish to make the journey, being well contented with the limited view from the valley. If the teacher stands at the top and merely beckons, few will heed him. Often he will show them pictures of views from the heights and play them fine music which they will increasingly be able to perceive more clearly. Naturally he will try to take them up by the most economical route. As far as possible he will choose a path which is attractive in itself. As Mursell (1948) remarks about developing musical responsiveness in playing the violin, the phrase *per aspera ad astra* no doubt applies. But if the pupil has hardly caught so much as a glimpse of the *astra*, the *aspera* are very likely to defeat him. If we start the pupil climbing an interminable range of arid hills, encourage him with tales of vast views and flowing streams far ahead and refuse on principle to give him so much as a drink of water, he is exceedingly apt to leave his bones by the trail.

Commenting on the Newsom Report (1966), the Standing Conference for Amateur Music (1966) stated 'Somehow, excellence is recognised by the most untutored of men and not even the cynical youngster fails to appreciate the quality of music when it is made by a superlative artist. No child should be allowed to go through his school life without coming into contact with the greatest musical works in our heritage.' A young teacher in a new school was once sent to a class which she found in uproar. By some inspiration she sat down at the piano and played *Jesu, Joy of man's desiring*. Gradually silence fell on the class till all were quietly listening.

One isolated experience of true aesthetic appreciation may not mean much to children who are obsessed with the more animal response to 'pop' and beat music. But it is important that they should *feel* the difference. They do not have to give up their 'pop', in order to enjoy more valuable types of music. The main point is to make them feel that the more worthwhile music offers much enjoyment, even if it requires some effort to understand. They must feel that the teacher is leading them not to 'his' land, still less to some alien land which the establishment decrees they *ought* to like, but to a land that will be 'theirs'.

If we adopt the service of the aesthetic and the beautiful as the primary aim of music teaching, it can be used as a criterion by which to judge our musical activities. Technical exercises are valuable only in so far as they ultimately enable the player to produce a fine performance. 'Creative' music is worthwhile if it is inspiring the children to produce something of aesthetic merit on however modest a scale.

Beauty, like happiness, cannot be pursued too directly. In everyday teaching one often has to be conscious of ways and means, rather than of fundamental aims. But it is important that the underlying aims should not be lost sight of.

XXVI

Aids to Learning

PROGRAMMED LEARNING AND TEACHING MACHINES

Reference was made in Chapter XVI to Skinner's use of teaching machines for training in rhythm and pitch discrimination.

According to Skinner's basic principle, what is rewarded (reinforced) will be learned, while other random actions will drop out. What the student needs in order to learn is a reward after each step. Long-term goals, such as the desire to become a doctor or a concert pianist, are all too often feeble motivation for current learning. To try to maintain the student's motivation school marks, blame or praise by the teacher are commonly used to bridge the gap between the distant goal and the present learning situation. But there is often some time lag between completing an exercise and finding out whether or not it is correct. Moreover, to increase the effect of this form of incentive, competition between pupils is often fostered. This means that the weaker pupil must inevitably suffer competitive failure.

Skinner (1961) believes, however, that human behaviour can be remarkably influenced by small results, such as come from a gain in competence, learning to manipulate the environment, or even simply moving forward after completing one stage of an activity. In his teaching machine programmes he aims at ensuring that a correct response will be made about 95% of the time. This means that the material must be divided into small steps and cues to the correct response incorporated. The amount of cueing information is usually reduced as learning proceeds till it is no longer necessary.

Programmed learning, whether presented by some type of 'machine' or in the form of a book, has three identifying characteristics:

1. The material is presented in a logical sequence and in step sizes appropriate to the task and the needs of the learner.

2. A response is elicited from the student to each of the steps (or frames, as they are usually called).

3. The student is given immediate knowledge of the results of his response.

The need to break down the material into small sequential steps in itself leads – or should lead – to considerable improvement in presentation, and is a salutary exercise for most teachers. While the good teacher is constantly reviewing in his own mind his aims and his means of achieving them, the discipline of trying to write a programme may make the teacher realise that he has been in the habit of leaving too much to the pupil.

Not everyone would agree that programmed materials should reflect Skinner's concept of using very small steps in order to avoid any error. Crowder (1960) believes that error can provide a means of diagnosing individual problems and thus become an important factor in the learning process. He developed a type of programming known as *branching*, in which the student's response controls the material he sees next. If he passes the test question he is automatically given the next unit of information and the next question. If he fails, the nature of his error is explained to him and he is provided with additional practice. In this branching technique the response to be made is not necessarily a constructed one. He may select his answer from a given set of possible answers; for in Crowder's view it is the choice he makes that is important. Both branching and linear (the Skinnerian) types of programme had their partisans in the earliest experiments. Current research seems to be showing that each has its uses.

Research with programmed learning has confirmed on the whole that it has a very valuable contribution to make to many educational problems.

Good results, however, do not always seem to depend on following the principles laid down by Skinner and Crowder too rigidly. For example, responding by thinking the answer without writing it down has sometimes been found to be just as effective as well as quicker than writing the answers (Leith and Burke, 1966). Again, though one of the great advantages claimed for programmed learning is that the individual can work at his own pace, better results are sometimes obtained when pupils work in pairs or in small groups (Amaria, Biran and Leith, 1966). It seems that the brighter pupil can help his duller classmate to attain a satisfactory standard without holding up his own progress.

PROGRAMMING IN MUSIC EDUCATION

So far only a few programmes have been developed in the field of music education although there is undoubtedly considerable scope for useful programmes. As Ihrke (1963) pointed out, the problem of providing adequate individual instruction in music is very great. Under present teaching conditions the most a student can expect is four–six hours a week of individual instruction. For the rest of the time he may listen to lectures and to music performed by himself or by others, but the feedback he gets is delayed and imprecise. The wide range of musical discriminations that competence requires just cannot be provided by conventional teaching methods. The use of programme learning in music could free the teacher from much of the tedium of individual drill and enable him to turn his attention to means whereby students with a repertoire of precise skills and concepts could employ these ideas in listening, performing and composing.

THEORY OF MUSIC

The first programmes to be published have been concerned with the rudiments, or what the Americans call the fundamentals, of music. This is understandable since a knowledge of notation is relatively easy to present in a series of logical steps.

Robert Barnes: *Fundamentals of Music* (1964). This was developed at the Ohio State University with the needs of prospective elementary school teachers in mind. In Barnes's early experiments he used a programmed text in conjunction with classroom instruction. His subjects were 42 undergraduates, enrolled in two classes. Both classes were taught the fundamentals of music by the same teacher. One class was given the programmed text as an auxiliary learning aid. At the beginning of the experiment a pre-test of 100 items was used to obtain information on the students' existing knowledge of rudiments. After five weeks of instruction, a post-test composed of the same questions presented in a different order was administered. A final test with the same items, presented in a different random order, was given five weeks after the post-test to see how much of the learning had been remembered. The experimental group did significantly better at both the post-tests than those who had not taken the programme.

The programme seeks to teach time signatures, note and rest values from the semibreve to the semiquaver, dotted notes, notation in both treble and bass clefs, and major and minor key signatures. Included are a chapter on the piano keyboard and one on solfa syllables. Most of Chapter 8 and Chapter 10 are review chapters. Each chapter, except 8 and 10, is preceded by criterion questions. If the student's answers to these are right, he is told that he may skip the chapter and go on to the next. The programme is presented in book form, with the answers printed in a shaded area on the left of the page. He works down the page and then on to the next page as in normal reading. The correct answer is to be written under the answer given if a mistake is made. If as many as 10% of the answers in any one chapter are wrong, Barnes advises the student to repeat the chapter. The normal time to complete the course is said to be 4 to 6 hours.

The programme seems to be enjoying some use by American teachers in their general music classes in secondary schools, as well as in colleges offering fundamentals of music courses. Barnes does not suggest that the student should try to play what he hears on the piano, but clearly envisages that the students using the programme will be receiving experiences of actual music from the class teacher or other sources.

Austin Andrews and Jeanne Wardian: *Introduction to Music Fundamentals* (1963). This programme is also aimed at providing a course for elementary school teachers, but includes 36 popular school songs, which the student learns to play with simple accompanying chords on the piano.

During the development of the programme, Wardian (1963) experimented with 56 subjects, half of whom used the programme, while the other half had lessons from a teacher. After an eight-day period of instruction, both groups were given a test covering the material. The group who had used the programme did as well as the control group even though they had spent a significantly smaller amount of time learning the materials. In later experiments, the programmed materials were used in conjunction with classroom teaching. Once again, the experimental subjects did significantly better than the control group at an examination after the training period.

The 480 frames are arranged three to the page. The student starts with frame 1 at the top of page 1, then turns to the top of page 3 to find the answer, which is printed on the left of frame 2. Having checked his answer and attempted frame 2, he turns to page 5 for the answer. After working to the end of the book he returns to page 1 to do the frames printed in the middle of the page. This device ensures that the student does not inadvertently see the answer before trying the question but it does entail a considerable amount of page turning. Besides the 36 songs, the student is from time to time urged to try on the piano what he is learning. For example, he is told to play the interval of the major third and listen to its distinctive sound. Andrews and Wardian quote results achieved with the programme in the sort of teaching situation for which it was intended. For example, a group of graduate students worked through the programme in two weeks; the average number of correct points out of 100 rose from 14 before the training to 97 after it.

John Clough: *Scales, Intervals, Keys and Triads: A Self-Instruction Program* (1964). Clough assumes that the student knows already the names of the notes in treble and bass clefs, and the names of the notes on the piano keyboard. His programme is therefore confined to four more advanced topics. Each of the four parts comprises several sets, each of which contains some 20–50 frames. The whole programme is likely to take the student 5 to 12 hours to complete. Clough suggests it might either be completed in two or three weeks, or else the sections might be studied separately. He recommends that sets found particularly difficult should be revised. He states that all sections have been tested by at least 150 students and some by more than 400 in the course of the development of the programme.

Paul Harder: *Basic Materials in Music Theory* (1965). This programme is organised under 11 different chapters, which include, besides the usual rudiments materials, information on church modes and the basic principles of acoustics.

The programme was developed at Michigan State University for use by college students. Like the three programmes mentioned above, it could also be used by secondary school pupils. Harder recommends the student to play or sing each item as it is presented in order to relate the symbol to the sound. This would of course mean that the student would have to work at a piano or at least in a room of his own.

Gary Martin: *Basic Concepts in Music* (1966). Developed at the University of Oregon, this programme uses a branching technique. The student's understanding of each concept is tested before he is asked to read an explanation of it. According to his answer to a criterion question, he is directed either to pass on to the next question, or is given a short résumé or a long explanation. Self-evaluation tests are given at the end of each chapter. Martin has included a wider range of topics than were dealt with in some of the earlier programmes. He includes notation of rhythm and melody; intervals and chords; major and minor scales and chords; and the basic structure of music. The programme would be particularly valuable to students who already have some knowledge of rudiments. They can review quickly the concepts with which they are familiar and concentrate on their weak points.

OTHER PROGRAMMES

For children younger than those catered for by the programmes mentioned above, two of the United States companies which publish teaching machine programmes have produced junior courses. 'A Self-Tutoring Course in the Fundamentals of Music' was published in 1960 by T. M. I. Grolier for children of 10 years upwards. 'Elementary Music Reading' was produced in 1962 by the General Program Teaching Corporation for pupils, aged nine to twelve. These programmes could be used in Britain, though there are some minor differences in terminology, for example, the Americans call the crotchet 'quarter note'.

Whenever possible, the learning of such musical symbols should go hand in hand with learning the sounds they represent, if only because the rudiments of music are not intrinsically very interesting. Though in some situations 'crash' courses may be needed, it would probably be better to spread a programme over a few weeks.

Experimental programmes covering more than a knowledge of rudiments have been devised at the Northwestern University and at West Virginia University. Theodore Ashford (1966) programmed the material in the first seven lessons of the Music Theory course offered to first-year students at the Northwestern University. His programme included the characteristics of tone and the basic principles of composition, as well as scales and intervals. During the three-week instruction period, 23 students took the programme while the same number attended a regular course. The amount of time the latter spent on homework was recorded, as was the number of hours spent on the programme by the experimental group. A significantly greater improvement over the pre-test scores was achieved by the experimental group. Moreover, the average number of hours they spent on the programme was 7·8, while the class time plus homework averaged 13·7 hours for the control group. When a third test was administered some eight weeks later there was no difference between the groups with regard to retention, nor was there any difference between their performance on the college examination. Ashford concluded that a combination of programmed instruction with class teaching might help to solve the problems caused by an increase in the amount of subject matter, which often has to be taught within an unchanged amount of teaching time.

At the West Virginia University School of Music, Leo Horacek (1963) produced an experimental programme which covered the entire content of the first-year music theory course, except for keyboard harmony. The programme consisted of a number of written lessons, but most of the students' time was spent with a tape through which they listened to recordings usually in conjunction with written materials. The course included discrimination and identification of intervals and chords; melodic and rhythmic dictation; sight-singing of melodies and intervals; and lessons involving short chorale-type harmonisations to which the student had to make varied types of response.

AURAL TRAINING

As long ago as 1949 Frank Cookson at Northwestern University obtained promising results from experiments with tape-recorded musical material for self-tutoring.

Ten years later Charles Spohn at the Ohio State University completed an investigation of the possibilities of using tape-recorded material to develop ability to discriminate intervals. His procedure resembled programmed instruction in that the material was presented in a series of small steps, but immediate feedback of results was not given. After training, the errors of his experimental group had decreased by 80% compared with the 58% achieved by a group taught by traditional methods.

With two colleagues, William Poland and Caroline Arnold, Spohn (1963) then programmed the task of learning to identify the ascending melodic intervals of the major scale. Each interval was sounded for four counts. After a pause for the student to write down his answer, the correct response was announced on the tape. Finally the interval was played again. The student checked whether his answer was right or wrong. Set 1 of the series consisted of the four intervals found from previous research to be the easiest (octave, minor second, major second and minor third). If 46 out of 48 responses were correct, the student was allowed to proceed to the next category. If the first performance was below the criterion, however, he had to repeat the four intervals till criterion level was achieved. He then had to work through another drill with the same intervals till 46 out of 48 responses were correct, before proceeding to the next level of difficulty. The training period lasted six weeks. 47 out of 77 first-year music students completed the whole series. The other 30 also improved just as much on the post-test in judgment of melodic intervals. But, whereas the more successful group showed equal improvement when the intervals were played harmonically, those who did not finish the course had gained less significantly.

Spohn has since extended his research to the investigation of rhythm, tone groups and melodies (Spohn and Poland, 1964; Spohn, 1965). In one experiment four groups were matched on the basis of a variety of music tests. Group 1 then worked through a programme in which they listened to the tape, and wrote down their answers. Group 2 were given notation and had to sing their responses which were recorded. Group 3 listened and then sang. With group 4, the material was presented visually only and the answers were written down. Each student worked on each of three programmes – intervals, rhythms and tone groups – for 10 weeks.

All four groups were tested after the training with the four methods of presenting the rhythmic groups, intervals and tone groups. As one would expect, in general the groups had gained most, compared with the pre-test, on the tasks for which they had been trained. However, there were three interesting exceptions. The group who had practised listening to intervals and then writing them did better at looking at the notes and then singing them. With rhythm, the group that had been given the names of the notes and asked to write them in rhythmic notation did better at hearing the rhythm and writing it down. Those whose training on tone groups had consisted of looking at the notes and singing them (i.e. sight-singing) did better on the tone group test which required them to listen and then sing the notes. This was presumably because when doing the task which they had practised, they had tried to sing the notes to themselves – which had been harder than when the notes were actually played. The most effective method for training in intervals seemed to be the one that required the student to listen and then to sing and name the interval. The best method of presenting rhythmic patterns was by aural stimuli requiring written response. This might be used as a preliminary activity before dictation and sight-singing. With tone groups singing after listening and giving a written response to the notation were easier than dictation and sight-singing.

Spohn, however, suspected that individual students differed in the method of training which they found most effective. In a later study (1965), Spohn experimented with two different methods, visual and aural, of presenting the correct answers. Group results showed that there was little significant difference between the two. Again, however, Spohn believed that individuals might differ in the method found most congenial.

Though undertaken to provide evidence on specific learning questions, Spohn's later research has also shown the effectiveness of programmed learning as a training instrument with aural materials. A programme on intervals is being published (Spohn, 1967).

Betty Kanable's (1965) main interest was in investigating the question – could sight-singing be improved by students working on their own with a programmed tape? From her own teaching experience she had observed that many students lacked an adequate background of theoretical knowledge and of aural training to enable them to take a full part in choral music. She developed a programme of 283 frames covering major and minor keys up to four sharps and flats. The majority of her melodies were four bars long. Four-track tape was used so that the student could hear the master tape record his answer and then play back the answer, followed by the correct version on the master tape. Because the student had to be able to match his own response with the master, Kanable confined the experiment to those who scored at least 7 out of 15 on an error detection test and who could match 4 out of 5 pitches.

Fifteen high school students at a Summer school course worked with a programme on 12 consecutive days with training sessions of 50 minutes each. Fifteen control students had similar training but with a class teacher. The majority of students were 15- or 16-year-olds and were typical of the students who would later attend college courses in music.

At the beginning and end of the experiment a sight-singing test on the material that had been studied was given. The students who had worked with the programme had improved slightly more than those who had attended the class. A few of the students commented that they had found the earlier items rather too easy and that the difficulty of the later items increased rather too steeply. But it sounds as if, with some further development, Kanable's programme could be very useful.

James C. Carlsen: *A Program for Self-Instruction* (1965). Carlsen's programme was the culmination of four years of research at the Northwestern University and at the University of Connecticut.

In his early experiments, Carlsen (1964) investigated whether melodic dictation could be learned as effectively from material programmed and presented on tape, as from class instruction. His subjects were students enrolled in a first-year ear-training class. Each member of the experimental group worked by himself with the programmed material. The control group was trained on similar melodies by traditional class methods. Results showed that the programmed technique was definitely more effective. This was particularly true with the more complex concepts. These were learned almost as well as the basic ones by programmed instruction, but the classroom teaching was much less effective with this more difficult material. Carlsen suggested, however, that the effects of carefully sequencing such complex aural concepts in a classroom situation ought to be further investigated. It certainly seems likely that a programme that permits the individual to control his own rate of progress, and to repeat items as often as necessary, would be especially useful with difficult material. One interesting point which emerged from Carlsen's experiment was that students in the experimental group who had high mathematical aptitude did better than those with lower aptitude. This suggested that individuals with mathematical talent are penalised when taught melodic dictation in a classroom situation.

Carlsen also compared a branching with a linear form of programme. The linear programme group used every frame of the material. The branching group used only selected frames, unless they made an error, in which case they branched for additional practice. No significant difference in achievement was found between the two groups. As the branching programme required more expensive two-track machines, the published version was presented in linear form. The student who requires additional practice is recommended to repeat any frames which cause him difficulty.

The published version consists of 570 frames. The student listens to each frame, records his answer and checks it against the answer printed at the right side of the page. Care is needed to mask the answers quickly so as not to see them inadvertently. The programme begins with perception of simple time and pitch elements, and progresses to eight-bar melodies in simple and compound time and in major and minor modes. Many items require the student to detect where the printed notation differs from the tune that he hears on the tape, and to write the correct version. To add variety and to provide experience of hearing the melodies played on different instruments, eight string and wind instruments are used besides the piano. The programme includes a self-analysis chart so that the student has a record of his weak points and can see which concepts need revision. Twelve tests are included as measures of how well each section has been learnt. They can also be used as criterion tests before the beginning of each section. Students scoring high can omit that section. In his Instructor's Manual, Carlsen quotes results obtained with two small groups which showed that their errors on a melodic dictation test decreased by about 40% after using the programme for five to nine weeks. A third group who worked with the programme over a period of six months and also had class practice in sight-singing and harmonic oral perception decreased their errors by 65%. Carlsen suggests that work periods of about 30 to 40 minutes are optimal, since the programme requires considerable concentration.

Carlsen's experimental results were obtained with students who worked completely without the help of a teacher. He believed that programmed learning in general might be more effective if incorporated as part of a course of instruction in which the teacher played an essential part.

The later frames reach the standard of the professional music student. A junior version for pupils with some knowledge of notation who are beginning to learn an instrument would also be very valuable. Meanwhile, students who are unable to progress as far as the later frames will find working through the earlier ones useful. The tapes are expensive. But for educational purposes, the cost is much less than a week's salary for even the most junior music teacher. With careful timetabling, Carlsen claims that one set of tapes is sufficient to enable 200 students to complete the course in a term.

Programmes on Harmonic Perception, Contrapuntal Perception, Detection in Ensemble Performance and Aural Perception of Structure and Style are promised for the future and should prove extremely useful.

PERFORMING SKILLS

Skinner's device required the child to tap in unison with or to echo a pattern of beats. The device both generated patterns and monitored the responses. An attempt to use actual musical examples is being developed by Walter Ihrke at the University of Connecticut. The training equipment consists of a printed manual read by the student, an electric organ, an electric rhythm-monitor and a stereo tape recorder capable of providing automatic tape stop. The student works at the organ keyboard tapping the items from the manual. An audible tape provides background music and metronome beats which set the tempo. The student's response is monitored by a second inaudible tape channel, which contains the magnetic signals which the student is trying to match. The signals from this tape and from the keyboard are transmitted to the rhythm-monitor, which electronically compares them. If his response is too early or too late, the student receives immediate feedback of the error by flashing lights. The tape recorder stops automatically at the end of each item. By means of a control panel, the student can repeat the previous item if he wishes – a feature which Ihrke's students found very reassuring.

Ihrke (1966) experimented with seven students selected at random from a class of 24, enrolled in a college course called 'Music for the Classroom Teacher'. The students were definitely not music oriented, though some had received sporadic training. Each of the seven spent two hours a week in the training lab until the full programme of 190 items had been completed, eight to ten hours in all. The musically untutored had very little more difficulty in the programme than those with some experience. All 24 students took a pre-test and a post-test which contained identical rhythmic material. The post-test was, however, much more difficult, being twice as long and played twice instead of three times. The number of errors was, therefore, likely to be considerably greater. The average percentage of errors made by the experimental group increased by only 12%, while that of the control group increased by no less than 235%. Proficiency in playing a rhythm had thus transferred to rhythmic dictation. It would be very likely to transfer to performance on other instruments. But in any case, as Carlsen (1965*a*) pointed out, with today's technology there is no reason why similar feedback devices could not be made in the form of instruments other than the keyboard.

All Ihrke's students enjoyed the automated training; five felt that they were definitely participating in the background music. One student was heard repeating an item several times, though she could play it correctly. 'I am repeating it because I am enjoying it so much,' she said when questioned. The value of this training as a preparation for ensemble playing is likely to be great. Ihrke has concentrated on rhythm as offering elements which are clear-cut and readily programmed, but foresees his methods being extended to other aspects of music.

Another possibility of providing feedback and reinforcing correct responses would be to use a keyboard that produced sound only when the correct key was depressed. Maltzmann (1964) has experimented with such an instrument. The task which he set his subjects required the matching of tones. He used a variety of procedures:

1. If the subject responded correctly to the sound he heard, the key he depressed produced the sound and a red light lit up. If he pressed the wrong key, no sound was produced and no light appeared.

2. Similar to (1) without the light.

3. A sound was produced, whether the key depressed was correct or not.

Procedures (1) and (2) where the subject heard only correct tones, proved to be more effective than method (3). Method (1) gave better results than method (2). Maltzmann (1965) has continued his experiments, studying for example the task of learning to discriminate different intervals. He found that once the third, the fifth, the seventh and the octave had been learnt in one key, the student could identify the same intervals in other keys. However, even after he had become nearly perfect in two keys practised separately, presenting intervals from two keys alternately greatly disturbed the performance.

OTHER AIDS TO LEARNING

There is a considerable need for an experimental approach to the evaluation of the uses of such equipment as the tape recorder in schools. It would seem, for example, to have considerable usefulness in allowing detailed analysis of performance whether by pupil, teacher or research worker to take place at leisure. But really carefully controlled studies are few.

One well-planned and executed study was carried out by Robert Biggs (1960) at the University of Iowa. Ten students who had volunteered to take part in the experiment followed a similar course of instruction in brass instruments as a controlled group, except that they used 70 minutes of a six-hour preparation period to listening to a tape-recorded lesson. The experiment continued for eight weeks. At the end Biggs was forced to conclude that there was no statistically significant difference in improvement in either performance, technique, or interpretation between the two groups. But we may wonder whether a longer project and larger numbers of students might not have shown some significant result.

Margaret Sears (1965) compared the effect of tape-recorded singing lessons with lessons from a class teacher. The recordings were prepared by a specialist music teacher and given to two classes of children aged six and seven over a period of nine weeks. The training consisted of tone matching and the singing of songs. Two other classes of similar age received lessons from the class teacher, who followed the tape-recorded lessons as closely as possible. The children attended well to the taped lessons and their singing improved more than those taught by the class teacher. Sears believed that tape-recorded lessons could have a useful place in schools, for example, in enabling more children to have the benefit of specialist teaching.

Another piece of apparatus that has enjoyed some experimental use with music materials is the tachistoscope. The tachistoscope is an apparatus for exposing visual stimuli for a fifth of a second or less. An early investigator, Charles Stokes (1944), tried to improve his subjects' visual span for musical notes by using the tachistoscope in 21 ten-minute lessons. Each lesson consisted of 40 slides of melodic material which was flashed on the screen before the students for $\frac{1}{10}$ second. As the training proceeded, the programme became increasingly more difficult, the horizontal span being increased from two notes to seven, and the vertical span from examples using a unison to those using a ninth. For the first five lessons students who had had musical training were asked to identify the example from a number of printed examples. After the sixth lesson the student had to judge whether an example played on the piano was the same or different from the example he had seen. Stokes believed that the students were showing improvement since the mean scores were staying constant in spite of the material becoming more difficult. However, on the criterion of performance on the Knuth achievement test, no significant difference between the experimental group and a controlled group was found.

More recently Harry Hammer (1963) has investigated whether the tachistoscope could be used as an aid to training in sight-singing. With two classes of ten-year-old children in Colorado whose training differed only in that the experimental group received 12 minutes of intensive drill where the material was presented by a tachistoscope as compared with the conventional method, the experimental groups improved more than the others. The tachistoscope training seemed to be more effective with less intelligent children than with the more musically talented children.

Believing that the principal cause of difficulty in reading music is inability to grasp rhythmic patterns, William Christ (1953) experimented with presenting rhythmic patterns by tachistoscope. Each pattern was first projected at speeds varying from $\frac{1}{100}$ second upwards, to a group of 11 music students at Indiana University. The students tried to tap the pattern, all tapping together. After each attempt the pattern was re-projected on time exposure so that faults could be corrected through immediate practice. This training yielded significant gains in improvement in ability to receive and reproduce rhythmic patterns, but Christ did not try to relate this drill to the reading of actual music. In all, the students spent ten hours on this practice. We may wonder what would have been the effect of spending a similar length of time on training in rhythmic patterns by some other method.

Heller (private communication) is engaged in experiments with a more sophisticated piece of electronic equipment. This apparatus 'will write out' on graph paper a readable picture of frequency and intensity changes of a musical melody *while it is actually being performed* by an instrumentalist or a singer. A pre-recorded graph of a musical phrase can be made by the teacher and used as a model for the student to emulate. Heller is finding that students do seem to be able to match the models with reasonable accuracy after several trials.

Experiments are currently proceeding at Stanford University to evaluate a computer-assisted teaching system which provides accurate information on the pitch of series of intervals sung by the student. The computer prints out whether each note has been sung correctly, or sharp, or flat; or if a different note from what was required has been sung, it names the pitch of the wrong note. It can also be programmed to add comments such as 'Congratulations' to encourage the successful student (see Kuhn and Allvin, 1967).

It seems reasonable to look forward to a time when students will be able to acquire basic skills more rapidly and when teachers will not have to supervise routine practice. However, much more research needs to be carried out to investigate under what conditions new equipment can best be used. Meanwhile it is perhaps comforting to find that traditional methods of instruction sometimes produce as good, if not better, results than methods requiring equipment.

William Graves (1963) compared the effectiveness of aural, visual, and 'conventional' methods of improving the intonation of 54 students aged 16 to 18. Each subject received one 30-minute private lesson every week for a term. Ten minutes of each lesson was devoted to work on intonation. The aural group tried to adjust their intonation by matching the pitch of reference tones sounded on an organ. The visual group used a Stroboconn. The 'conventional' method required pitch inaccuracies to be detected by the teacher and the pupils. The music used consisted of a major scale, a chromatic scale and tonic and dominant seventh chords. None of the methods was consistently better than the others, but the conventional method was generally the most effective method. The results did not depend on which instrument was being studied. Graves concluded that the development of good intonation seemed to depend on the quality of teaching, rather than on specific techniques or equipment.

CONCLUSIONS

Although only a few useful programmes have so far been developed, the possibilities for the future are immense. As Carlsen (1965*a*) suggests, aural and visual materials might be combined with a performance reading device in the teaching of the rudiments of music, of harmony and counterpoint, the history of music and form and analysis. (Carl Nelson (1967) is currently experimenting with programmes to teach form at Cortland College, N.Y.) Orchestration might be taught much more effectively with a programme illustrating the ranges and timbre of the instruments, and the possibilities and problems of combining them. Again, the training of conductors might be improved, if it included instruction in the analysis of errors made in ensemble performance. The material which Mueller (1956) envisaged as a means of testing the appreciation of compositions should lend itself to programming for training in the intellectual processes required to perceive the formal structure of music. Could we use programmed techniques to promote the aesthetic experience of music (which we have taken to be a fundamental aim of music education)? Maybe this is an area that would be better left to the teacher. Yet, why should there not be moments in the programme when the student is invited just to listen to how Beethoven has used the dotted quaver and semiquaver time pattern in his Minuet in G?

XXVII

Conclusions

So long as resources of specialist music teachers and of musical instruments for loan remain insufficient it seems sensible to give priority to the talented, but this should not mean rigidly dividing the gifted from the ungifted. The highly musical will find more scope for their talents as performers for there is a musical well-educated population of listeners. As we said in Part Four, every individual with only a modest share of musical ability can enjoy the greatest aesthetic experience of music. It is perhaps in fostering a love of beauty that music has special significance in the education of the average child.

The development of new teaching techniques may eventually enable the benefits of a musical education to be extended more widely. However exciting the possibilities opened up by new gadgets and programming techniques, what is important is how they are used. It would seem vital that all who are interested in teaching and education problems should be alert to the possibilities that such devices offer. But prompt evaluation of new techniques is required to see whether they are in fact better than older methods, or under what conditions they can best be used. This should help to ensure that they are not adopted as gimmicks for brief spells, only to be dropped because expectations that were unrealistic in the first place have not been valid.

RESEARCH POSSIBILITIES

The most sophisticated investigations are likely to be carried out by professional research workers. Yet much valuable research could be done by school or college teachers of music, who have the advantage of being in close touch with everyday problems. The spare time of musicians is usually well filled by conducting, ensemble playing or choir practice. However, those who feel an inclination to undertake research should not feel put off by the mystique with which it is sometimes surrounded.

Guidance on the design and planning of the project is vital to ensure that it is scientifically sound and feasible in practice. The actual carrying out of an intelligently planned investigation of limited scope is well within the reach of any teacher with patience and a genuine interest. Most standardised tests of musical ability can be administered by anyone who can work a tape recorder or record-player. The scoring is no more difficult than marking a rudiments of music exercise.

Guidance with the design of the experiment should include advice on appropriate statistical treatment of the results. The more elementary statistics are easy to understand (see, for example, Whybrew, 1962, Chapter 3). The actual calculations can be carried out by anyone who can add, multiply and use a table of square roots. The increasing availability of computers will take away the tediousness of lengthy calculations. The interpretation of experimental results requires insight, objectivity and fine judgment – qualities that are exercised every day by teachers in the course of their profession. Again, age is no bar to research. In fact many teachers in their 40s or 50s might find research a rewarding and rejuvenating experience.

For research to be of practical benefit in education, it is vitally important that music advisers, head teachers and music teachers should take an interest in the results and consider how they might be applied. An interest in research is, of course, only one of the qualities one might desire to find in a teacher. Many teachers produce excellent results without being at all interested in the findings of psychological investigations. But even they might profit from knowing on what sort of problem the expert is likely to be able to offer help.

THE ROLE OF THE TEACHER IN THE FUTURE

There does seem to be some foundation for hoping that the intelligent use of modern technological aids might banish some of the drearier tasks of the teacher. This does not mean that teachers could be superseded, but that they should have more time for the more interesting part of their profession.

According to Dr Murray Tondow of Stanford University, the role of the teacher in the future may be divided between two kinds of specialist. The curriculum specialist will devise the programmes and select those most suited to the current needs of the individual pupil.

The other type of teacher will work with students in small groups. He will discuss with them what the facts they have been learning mean, and how they are inter-related. He will thus assume the role that Plato envisaged for his philosophers.

XXVIII

Summary and Conclusions

'The search continues' . . . is perhaps the phrase most apt to sum up this book. It has been at best a progress report, drawing attention to what has been achieved, and even more important, suggesting lines of enquiry which might be followed up in the future.

Firstly, satisfactory means of assessing musical ability both for educational and research purposes have been evolved. Current tests cover a reasonably wide range of the various aspects of musical ability, from discrimination of fine differences of pitch to melodic memory and to the appreciation of subtle changes in style of playing or of music content. This last is a most difficult aspect to test by objective (or any other) method and present tests leave much to be desired. The age range for which useful tests exist now extends from eight to adult level. Such tests as the Wing can be used with confidence – so long as the confidence is tempered with good sense. High scores are indicative of a promising level of talent; low ones may often be partly due to misunderstanding of instructions or to some distracting or upsetting circumstances and should be treated with caution. In any case, as Whybrew (1962) remarked, results of tests should be regarded as tentative and supplemented by information from other sources.

Where teaching resources are limited, tests may provide a ready means of picking out those children most likely to profit from special opportunities to learn an instrument. But as far as practicable, tests should be used to suggest the type of training most likely to meet the child's individual requirements, thus helping him to express such gifts as he may have. The arousing and sustaining of interest is of fundamental importance for the full development of musical talent. A constant responsibility for both parent and teacher, indeed for us all, is the provision of a healthy social climate in which music is highly valued.

How far the full growth and flowering of musical talent depends on nature and how much on nurture is not known. The earliest years of life can be regarded as vitally important; whether they are 'critically' so is not yet clear. If parents can provide their children with opportunities to hear music and, if possible, to make music, but above all to enjoy music they can feel confident that they are nurturing whatever talent their children may have. If one or both of the parents can sing or play so much the better. Echoing back the child's own attempt to use his voice and later singing with him or encouraging him to sing with the radio or a record are valuable, and hardly require the vocal powers of a Joan Sutherland. With the child of school age the parents still have a vital role, besides paying for lessons and providing an instrument and a place to practice. Encouraging the child to play alone and in groups while not forcing him, taking an interest yet not intruding when he feels self-conscious, these are all familiar, but necessary, parental responsibilities.

However favourable the environment, genetic factors may still set a limit to the speed with which the individual learns to perform musical tasks and to the ultimate level of his achievement in music. Many, possibly most, people never develop their powers to the full. It is difficult to judge how near a person has come to reaching the limit of his capacity. But if a child from a musical home has been competently taught over a period of years and has studied diligently, he may reasonably be said to have enjoyed good opportunities of developing his potentialities. If his achievements remain modest compared with those of children with similar or lesser opportunities, we may suspect the operation of genetic factors.

As we have seen in Part Four, opinions differ about the best way of describing the complex structure of musical ability. Some psychologists and musicians believe that one important general factor underlies the performance of a comprehensive set of musical ability tests. Others stress the distinctions between the various aspects of musical ability, while admitting that they often overlap. Statistical studies of this question have so far been limited either because comprehensive tests have been applied to rather few subjects or because when a large number of subjects has been used, the number of tests has been small. There is still much that might be learned, for example, from comparing persons differing in age or in experience of music. On present evidence, it seems likely that a broad factor of ability to master music will be found as well as narrower and more specific ones.

Though a reasonably good level of general intellectual ability is required for successful achievement in music, most of the empirical evidence suggested that musical aptitude may be classed as a special ability.

Music can make a valuable contribution to the intellectual and emotional development of the child. Every child should be given the opportunity of feeling its power as a source of aesthetic experience, though some will eventually decide that other arts are more congenial and rewarding.

The development of new techniques in education hold promise of freeing the teacher from some of his more routine tasks. But how such techniques can best be assimilated into the music education programme requires careful appraisal.

Many questions, then, remain to be answered by research in the future.

One thing, at least, is certain. There are vast resources of musical talent which await development. The discovery and fostering of this talent at the earliest stage depends perhaps most of all on the co-operation of parents and teachers, calling on the technical advice which experts are always only too ready to give.

Appendixes

APPENDIX I

Description of Tests

MAINWARING TESTS OF MUSICAL ABILITY

Three Tests Pitch, Rhythm and Recall.
Details Published in Brit. J. Educ. Psych. (1931), **1**.
Pitch a. 16 items. Two notes are played. Are they the same or different?
20 items. Which two of three or four notes are the same?
b. Concept of 'high' or 'low'.
9 items. Do three notes go up or down?
5 items. Which of two notes is the higher?
20 items. In which pair of intervals are the two notes farther apart?
Rhythm 25 items presented with a metronome, or pencil tapping, or buzzer, or with rhythmic word groups. In each case the subject has to decide whether the metre is in two's, three's or four's.
Immediate Recall 5 tunes, 4–11 notes long, are played. 10 seconds after the playing of each, various questions are asked, e.g. did the last two notes go up or down or were they the same? Was it in two, three or four beats? The ten items are repeated.
Deferred Recall The subjects are asked three questions about *God Save the King* and three on *While Shepherds Watched.*

Reliability
Pitch ·81 (Mainwaring); ·77 (Fieldhouse, 1937).
Rhythm ·74 (Mainwaring); ·62 (Fieldhouse, 1937).

MADISON MUSIC TESTS

Two Tests Interval Discrimination, Tonal Imagery. Interval Test described in *Arch. Psychol.* (1942), **206**, 1–99.
Interval Discrim. Items.
Tonal Imagery 36 items. Four harmonic intervals are played at different pitch levels. One of the four is always different. Which?

Reliability
Intervals ·74; ·76; ·84.
Tonal Imagery ·76; ·84.

Validity

Intervals ·46–·72 (with music students); ·39–·71 (secondary school pupils).

Tonal Imagery ·24–·51; median ·41 (Christy, 1956).

SEASHORE MEASURES OF MUSICAL TALENTS

1919 VERSION

Six Tests Pitch, Intensity, Consonance, Tonal Memory, Time, Rhythm (added five years later).

Reliability

	(1) Range	(1) Median	(2) Range	(2) Median
Pitch	·51–·84	·71	·58–·90	·77
Intensity	·50–·88	·72	·55–·94	·75
Time	·41–·81	·58	·45–·62	·56
Consonance	·30–·62	·49	·35–·68	·46
Tonal Memory	·59–·94	·83	·66–·90	·77
Rhythm	·29–·68	·45	·30–·50	·45

(1) From Lundin (1967).

(2) From Farnsworth (1931), studies not included by Lundin.

Validity compared with music grades and teachers' ratings

	Range	Median
Pitch	·01–·60	·23
Intensity	·02–·49	·13
Time	−·14–·36	·17
Consonance	−·27–·41	·05
Tonal Memory	·05–·65	·30
Rhythm	−·15–·47	·19
Total	−·15–·73	·27

1960 EDITION (similar to the 1939 revision)

Six Tests Pitch, Loudness, Rhythm, Time, Timbre, Tonal Memory.

Ages 10 to Adult.

Time to Administer About one hour.

Published by The Psychological Corporation.

A more difficult form of the test, 'B' form, for use with music students was published in 1939, but has since been withdrawn.

Pitch 50 pairs of tones. Frequency differences from 17 to 2 cps. Is second tone higher or lower than the first?

Loudness 50 pairs of tones. Intensity difference from 4·0 to 0·5 decibels.

Is second tone stronger or weaker than the first?

Rhythm 30 pairs of rhythmic patterns. Are they the same or different?

Time 50 pairs of tones. Duration differences from ·30 to ·05 seconds. Is second tone longer or shorter than the first?

Timbre 50 pairs of tones. Each tone made up of fundamental and first five harmonies, the intensities of third and fourth being varied. Are the two tones same or different?

Tonal Memory 30 pairs of tonal sequences, 10 items each of three-, four- and five-tones. Which note is different?

Norms Percentile for each test separately, none for total score. Grades 4 to 5, 6 to 8, and Adult. Based on approx. 3,500 for Pitch, Rhythm and Tonal Memory (grades 4 to 5), on 2,500 for Pitch, Rhythm and Tonal Memory (grades 6 to 8), much smaller numbers for the other tests for these grades (over 4,000 for all tests at Adult level).

Reliability
(*From test manual*)

Pitch	·82–·84
Loudness	·74–·85
Rhythm	·64–·69
Time	·63–·72
Timbre	·55–·68
Tonal Memory	·81–·84

Validity Questionable, except for Pitch, Rhythm and Tonal Memory.

Subjects:	Pitch	Rhythm	Tonal Memory	Total	
10-year-olds					
Success at violin	·33	·33	·41		(Manor, 1950)
Success at clarinet	·09	·00	·06		
Success at trombone	·14	·14	·15		
Music students					
'Musicality'				·46	(Kyme, 1956)
Performance	·13–·15	·18–·31	·14–·27	·34–·42	(Christy, 1956)
Theory and composition	·12–·19	·19–·46	·27–·41	·34–·49	
Theory grades	·30	·15	·12		(White, 1961)
Theory grades	Zero correlations with all tests and total				(Roby, 1962)
291 children (10 to 16 yrs)					
'Musicality'	·11–·45	·19–·27	·36–·50		(Rainbow, 1965)

KWALWASSER – DYKEMA MUSIC TESTS

Ten Tests Pitch, Quality, Intensity, Tonal Movement, Time, Rhythm, Tonal Memory, Melodic Taste, Pitch Imagery and Rhythm Imagery.
Ages 10–Adult.
Time to Administer One hour.
Published by Carl Fischer Inc.
Pitch 40 items. Does pitch of each tone remain the same (S) or does it rise or fall (D)?
Quality 30 items. Two notes are played twice. Is second on same (S) or different (D) instrument?
Intensity 30 pairs of tones or chords. Is second louder or softer than first?
Tonal Movement 30 four-note Tonal Patterns requiring completion. Should a fifth tone be above or below fourth?
Time 25 items of three tones each, first and third are of equal lengths. Is second same length as first or third or different?
Rhythm 25 pairs of Rhythmic Patterns that differ in intensity or duration, or in both. Is second (S) or (D)?
Tonal Memory 25 pairs of Tonal Patterns from four to nine tones long. Is second (S) or (D)?
Melodic Taste 10 pairs of two phrased melodies. First phrases are the same, the second different. Which second phrase makes the more appropriate ending?
Pitch Imagery 25 tonal patterns in notation. Are these (S) or (D) from those played on a record?
Rhythm Imagery 25 rhythmic patterns in notation. (S) or (D) from those heard on the record?
Norms Percentile norms for each test and for total scores for grades 4–6, 7–9, and senior high school. Based on 'thousands of scores'.

Reliability No information given in the test manual. According to independent studies much lower than comparable tests of the Seashore battery.

	Range	Median	
Pitch	–·05–·63	·34	
Quality	·10–·66	·36	
Intensity	–·10–·60	·15	
Tonal Movement	·37–·85	·68	
Time	·00–·63	·33	
Rhythm	·04–·48	·29	
Tonal Memory	·43–·73	·55	
Melodic Taste	·06–·61	·35	
Pitch Imagery	·14–·45	·33	Whybrew (1962)
Rhythm Imagery	·20–·40	·31	after Lundin

Holmes version (based on 237 students, aged 15–18)

Pitch	·72	Tonal Memory	·73
Rhythm	·71	Tonal Movement	·88
Time	·50	Melodic Taste	·43
Intensity	·79	Total	·91
Quality	·70		

Validity Doubtful except for discriminating most musical from least musical of a group.
The comparisons with teachers' ratings and grades.

	Range	Median	
Pitch	−·18–·23	·00	
Quality	−·10–·21	·14	
Intensity	−·11–·29	·13	
Tonal Movement	·00–·31	·18	
Time	−·13–·27	·01	
Rhythm	−·04–·31	·16	
Tonal Memory	·02–·45	·26	
Melodic Taste	−·19–·31	·01	
Pitch Imagery	·00–·59	·31	Whybrew (1962)
Rhythm Imagery	·01–·46	·29	after Lundin

KWALWASSER MUSIC TALENT TEST

Two forms A for ages 13 and over; B for ages 10 to 12. Each form gives scores on Pitch, Time, Rhythm and Loudness.
Time to Administer 10 minutes each form.
Published by Mills Music Co., New York.
Form A 50 pairs of short melodic patterns. Second differs from first in Pitch, Time, Rhythm or Loudness. Choice of two answers given for each item.
Form B Similar to A, but only 40 items, and easier.
Norms Form B, norms for grades 4 to 6; Form A, for junior, and for senior, high school levels.

Reliability Not mentioned in manual.
·48 (Farnsworth, 1959); ·59 (Bentley, 1955).

Validity Not mentioned in manual.
·46 with music grades (Bentley, 1955).
·34; ·34; ·39 with teachers' rating (Petzold, 1960).

THE DRAKE MUSICAL APTITUDE TESTS

Two Tests Musical Memory two equivalent forms (A) and (B), Rhythm two forms (B) more difficult than (A).

Time to Administer About 20 minutes for each form of each test.

Ages Eight years to superior musical adult.

Published by Science Research Assoc.

Memory 54 items – 12 melodies each played from 2–7 times. Is each repetition same as original or has key, time or notes been altered?

Rhythm 50 items. Subject has to continue to count a beat established by a metronome, during silence till told to stop. Number recorded is compared with correct answer. In (B) form counting is done against a distracting beat.

Norms Memory percentile norms for two-yearly intervals from 7–22 for non-music students, i.e. with less than five years of musical training, based on a total of over 4,300 cases; and for three-yearly intervals between 11 and 23+ for music students based on 1,400 cases.

Rhythm One set for all ages, but separate for music students based on approximately 1,300 non-music students and nearly 350 music students.

Reliability
(*From test manual*)

	Range	Median
Memory (A) + (B) forms	·85–·93	
Rhythm, Form A	·56–·95	·86
Form B	·69–·96	·775

Validity From manual compared with teachers' ratings.

	Range	Median
Memory	·32–·91	·55
Rhythm		
Form A	·31–·82	·58
Form B	·41–·83	·67
A + B	·31–·35	·58
Memory	·09–·50	·42 (Lundin, 1949)
	·17–·32	·24 (Christy, 1959)

OREGON MUSIC DISCRIMINATION TEST

No longer commercially available.

INDIANA–OREGON MUSIC TEST

Standardisation proceeding. See p. 34.

WING STANDARDISED TESTS OF MUSICAL INTELLIGENCE

Seven Tests Chord Analysis, Pitch Change, Memory, Rhythm, Harmony, Intensity, Phrasing.

Ages Eight to Adult.

Time to Administer One hour.

Published by National Foundation for Educational Research.

Chord Analysis 20 items. How many notes in the chord?

Pitch Change 30 items. Have the two chords been repeated exactly, or has note moved up, or down?

Memory 30 pairs of tunes from three to ten notes long. Which note has been changed on the second playing?

Rhythm 14 pairs of tunes. Is second the same as the first? If different, which is the better version?

Harmony; Intensity; Phrasing Similar to Rhythm, except that harmonisation, intensity or phrasing may have been altered.

Norms In five grades for total scores and tests 1 to 3, from 8 to 17 (Adult); based on nearly 10,000 cases. Scores can also be converted into Musical Quotients.

Reliability

·91 (Whole test)
·89 (Tests 1–3)
·84 (Tests 4–7)
·90 (Wing, 1962)
·86 (Whole test) (Bentley, 1955)
·90 (Whole test) (subtests ·65 to ·85) (Buros, 1960)
·80; ·82 (Whole test)
·78; ·86; ·89 (Tests 1–3)
·28; ·42; ·50 (Tests 4–7) (Heller, 1962)

Validity Good

With teachers' ratings: ·64 to ·90 (Wing, 1948)
·83 (Cain, 1960)

Significant differences found between 'above average', 'Average' and 'below average' RMSM junior musicians for 127 out of the 136 test items (Newton, 1959).

Significant differences between actively musical and unmusical groups (Whittington, 1957).

All members of National Youth Orchestra and all, except one, professional music students at Eastman School of Music achieved grade 'A'.

GASTON TEST OF MUSICALITY

17 Interest in Music Items
22 Tonal Items
Ages 10–18.
Time to Administer 40 minutes.
Published by Odells Instrumental Service, Kansas.

Tonal Items

5 Items Subject has to find a given note in a chord.

5 Items Melody of 4–8 bars answer sheet has to be compared with melody played for possible difference in note or rhythm.

5 Items Should final note be higher or lower than last one played?

7 Items Melodic Memory. Is second version same or different from first?

Separate percentile norms for interest and for aptitude. Separate norms for girls and for boys at Grades 4 to 8 and 9 to 12 for Interest.
Separate norms for boys and for girls at five levels (Grade 4 to Grade 12) for Aptitude. Norms based on a total of nearly 6,000 cases.

Reliability

From manual	Grades 4 to 6 and 7–9	·88
	Grades 10–12	·90
	Grade 12	·84 (Bentley, 1955)

Validity From manual – association between teachers' ratings and scores significant at ·05% level only for Grades 10–12 and 4–12.

Items 19–33 Too easy for Bentley's subjects but Melodic Memory items most discriminating of all tests investigated in distinguishing instrument from non-instrument playing groups.

WHISTLER & THORPE MUSICAL APTITUDE TEST

Five Tests Rhythm Recognition, Pitch Recognition, Melody Recognition, Pitch Discrimination and advanced Rhythmic Recognition.
Age 10–16.
Time to Administer 40 minutes.
Piano Version published by California Bureau, 1950 – no recording available.

Rhythm Recognition 10 pairs of items. Is second same as or different from first?

Pitch Recognition 10 items. A single tone is played followed by a 4-bar melody. How many times does the tone appear in the melody?

Melody Recognition 25 pairs of melodic patterns. Is second same as or different from first?

Pitch Discrimination 15 pairs. Is second chord same as first? If not, is it higher or lower?

Advanced Rhythm Recognition 15 pairs of items similar to first test with slightly intricate rhythms.

Percentile norms for each grade from 4–8 and for Grades 9 and 10.

Separate norms for rhythm, pitch and melody and the total scores.

Reliability Range from ·64 (rhythm)–·87 total scores; From manual ·745 (Bentley, 1955).

Validity compared with teachers' ratings and participation in musical groups range from ·19–·56; ·19 (Kyme, 1956).

THE GORDON MUSICAL APTITUDE PROFILE

Three parts Tonal Imagery (Melody and Harmony).
Rhythm Imagery (Tempo and Metre).
Musical Sensitivity (Phrasing, Balance and Style).

Melody 40 items played on violin. Tune and answer. Is answer a melodic variation of tune or is it different?

Harmony 40 tune and answer items played on the cello, upper part played by violin remains the same. Is answer a melodic variation of tune or is it different?

Tempo 40 items. Is end of answer at same or different tempo than end of tune?

Metre 40 items. Is there a change of metre e.g. from duple to triple at end of answer?

Phrasing 30 pairs of items. Which is performed with the better musical phrasing?

Balance 30 pairs of items. Which of the pair has better ending?

Style 30 pairs of items. Which is played in the better style?

Percentile norms based on nearly 13,000 pupils from 18 of the American States, for Grades 4–12, for each subtest, for each of the three parts, and for total scores. Separate norms for musically select students at three levels.

Reliability

	1 (From manual based on all students in the standardisation sample.)	2 (Over 1,000 pupils aged 10–18 (Tarrell, 1965).)
Mel.	·73–·85	·67–·80
Harm.	·66–·85	·67–·83
Ton. Imag. (Mel. + Harm.)	·80–·92	·80–·89
Tempo	·72–·85	·60–·82
Metre	·66–·85	·60–·84
Rthm. Imag. (Temp. + Metre)	·82–·91	·78–·86
Phrasing	·67–·78	·60–·72
Balance	·66–·79	·60–·89
Style	·66–·80	·60–·74
Sensitivity (Phr. + Bal. + Stl.)	·84–·90	·70–·84
Total	·90–·96	·86–·93

Validity

	Teachers' ratings		
	1 From manual based on 400 students		2 Culver (1965)
	Range	Median	
Mel.	·37–·88	·52	
Harm.	·52–·72	·64	
Ton. Imag.	·54–·83	·67	
Tempo	·48–·66	·58	
Metre	·57–·71	·65	
Ryth. Imag.	·64–·74	·66	
Phrasing	·19–·66	·36	
Balance	·20–·66	·48	
Style	·44–·87	·57	
Sensitiv.	·48–·85	·60	·53; ·72
Total	·64–·97	·79	·72
Ton. + Rth. Imag.			·69; ·80

Compared with performance of selected test pieces

	3 Tarrell, 900 pupils (1965) Range	Median	4 Fosha (1964) Range	Median
Ton. Imag.	·25–·43	·380		
Rhyth. Imag.	·13–·41	·285		
Sensitiv.	·17–·28	·235		
Total	·24–·43	·380	·12–·55	·290

See also page 38 and page 182.

BENTLEY MEASURES OF MUSICAL ABILITY

Four Tests Pitch Discrimination, Tonal Memory, Chord Analysis and Rhythmic Memory.

Age 7 or 8–14.

Time to Administer 20 minutes.

Published by Harrap.

Pitch Discrimination 20 items ranging from 26 cps to 3 cps. Is second higher, lower or same as first?

Tonal Memory 10 pairs of 5 note tunes. Is second same as first? If different, which note has been changed?

Chord Analysis 10 items. How many notes in the chord?

Rhythmic Memory 10 pairs of Time Patterns. Is second same as first or if different which note has been changed?

Norms divided into five grades, for ages 8–14. Based on testing some 2,000 children.

Reliability

Pitch	·74
Ton. Mem.	·53
Chord. Anal.	·71
Rhyth. Mem.	·57
Total	·84

Validity Significant association between test scores and teachers' estimates of the musical ability of three groups of children. Four groups of musicians or music students all made high scores.

LUNDIN MUSICAL ABILITY TESTS

Five Tests Interval Discrimination, Melodic Transposition, Mode Discrimination, Melodic Sequences, Rhythmic Sequences.

Unpublished

Intervals 50 pairs of items. Is second interval same or different?

Melodic Transposition 30 pairs of melodies. Second playing always in a different key but one or more notes may be altered. If transposed back to original key would it be the same or different?

Mode Discrimination 30 pairs of chords. Are both either major or minor (same) or is one major and the other minor (different)?

Melodic Sequences 30 items each with four melodic patterns. Has fourth pattern been changed?

Rhythmic Sequences Similar to Melodic.

Reliability

	167 Music Students	196 Unselected Students
Interval	·79	·71
Mel. Trans.	·65	·71
Mode Disc.	·65	·10
Mel. Seq.	·70	·77
Rhythmic Seq.	·60	·72
Total	·89	·85

Validity Compared with teachers' ratings on Melodic and Harmonic Dictation, written harmonisation performance and general ability in music.

	Range	Total of Ratings
Intervals	·32–·66	·48
Mel. Trans.	·26–·52	·45
Mode Disc.	·35–·51	·49
Mel. Seq.	·38–·56	·57
Rhythmic Seq.	·10–·33	·26
Total	·43–·70	·69

Very significant difference found between mean scores of 60 music students and 100 unselected students for each test and for total.

THACKRAY TESTS OF RHYTHMIC APTITUDE

Tests of Rhythmic Perception

Seven Tests Counting, Tempo, Duration of Sounds, Duration of Silence, Accent, Comparison of Rhythms and Rhythmic Patterns.

Unpublished

Counting 20 items. How many separate sounds in each item?

Tempo 10 pairs of eight sounds. Is second same in tempo? If not, which is quicker?

Duration of Sounds 10 pairs of sounds. Is second same length? If not, which is longer?

Duration of Silence 10 pairs of two sounds. Is interval of time between the two the same? If not, which is the longer?

Accent 10 items of eight sounds in which one or more of the sounds is stressed.

Comparison of Rhythms 10 pairs. Is second the same? If not, which has the more notes?

Rhythmic Patterns 10 items each given 3 or 4 times without a break. How many sounds are in each pattern?

Rhythmic Performance

Five Tests Tapping different rhythmic groupings with correct number of taps. Maintaining a steady tempo of regular pulses given at different speeds. Reproduction of different time patterns with correct values of notes and rests. Reproduction of eight beats with accents in the correct places. Reproduction of rhythm of short melodies, with correct tempo, time pattern, accent and tone gradations.

ALIFERIS MUSIC ACHIEVEMENT TEST (1954) – COLLEGE ENTRANT LEVEL

Three sections Melody, Harmony and Rhythm as described on page 43.
Time to Administer 40 minutes.
Published by University of Minnesota Press.

Reliability

From manual	Melodic section	·84
	Harmonic	·72
	Rhythmic	·67
	Total	·88

Validity From manual compared with music grades

	Test Manual		White (1961)	Roby (1963)
Melodic	·54	·57		·64
Harmonic	·41	·53		·66
Rhythmic	·46	·25		·37
Total	·61	·53	·63	·73
Melodic + Harmonic				·77

THE ALIFERIS MUSICAL ACHIEVEMENT TEST – COLLEGE MIDPOINT LEVEL (1962)

Three sections Harmonic Elements, Melodic and Rhythmic Idioms.

Reliability

Harmonic	·90
Melodic	·84
Rhythmic	·89
Total	·92

Validity Compared with Music Grades Range from ·40–·51.

FARNUM MUSIC NOTATION TEST

40 Melodic phrases.
One of four bars of each Melody as played is different from the melody in notation. Which?

Reliability Range, ·78–·91 (from manual); ·89 (Bentley, 1955).

Validity Compared with Watkins–Farnum performance scale, ·40–·61; compared with Music Grades, ·49 (Bentley, 1955).

APPENDIX II

Factorial Studies of Music Tests

Introduction

In the early years of the century the eminent English psychologist, Charles Spearman, observed that when different tests of ability are correlated the correlations are usually positive. They therefore seemed to be measuring some common factor. He developed a method of factor analysis to determine the existence of such general factors and to aid their identification. A common or general factor has positive loadings in all the tests. Group factors are obtained when only certain tests have positive loadings, the others being zero or negative.

The principal object of factor analysis is to simplify the description of the data by reducing the number of necessary variables, or 'dimensions'. If 20 tests were given to 100 persons, each individual's performance would be described along 20 dimensions, corresponding to the scores on each of the 20 tests. If by factor analysis five factors were found to be sufficient to account for all the common variance covered by these 20 tests, these five new dimensions could be substituted for the original 20 in describing each individual. It should then be possible to construct tests measuring each of these dimensions, or alternatively, to choose from among the original tests those that provide the best measures of the final factors.

All techniques of factor analysis begin with a complete table of intercorrelations among a set of tests. This is known as a correlation matrix. Every factor analysis ends with a factor matrix – a table showing the weight or loading of each of the factors in each test.

An understanding of the results of factor analysis calls for psychological insight rather than statistical training. The nature of a particular factor can be seen from examining the tests having high loadings on that factor and trying to discover what psychological processes they have in common. The more tests there are with high loadings on a given factor, the more clearly can the nature of the factor be defined. Factor loadings are expressed on the same scale as correlation coefficients: from $-1{\cdot}00$ through zero to $+1{\cdot}00$. In fact, factor loadings can be regarded as correlations of each test with each factor (or with what is common to a group of tests). Very low loadings can be ignored since they represent only chance fluctuations from zero, and are of little help in identifying factors.

It is customary to represent factors geometrically as *reference axes* in terms of which each test can be plotted. The position of the reference axes is not fixed by the data. The original correlation matrix determines only the position of the tests *in relation to each other*. Factor analysts can therefore rotate the axes till they obtain the most satisfactory and easily interpretable patterns. Such a procedure is analogous to measuring longitude from New York instead of from Greenwich. Thurstone, the American psychologist, recommended two criteria for rotation: 'positive manifold' and 'simple structure'. 'Positive manifold' requires the rotation of axes to such a position as to eliminate all significant negative loadings. Most psychologists regard negative loadings as inapplicable to aptitude tests, since they imply that the higher the individual rates in the particular factor the poorer will be his performance on that test, e.g. the better he is at arithmetic, the worse he will be at French. The criterion of simple structure means that each test shall have loadings on as few factors as possible. The aim of these criteria is to obtain factors that can be readily and unambiguously interpreted. In practice, however, results are rarely clear-cut and much room is left for subjective interpretation – or misinterpretation.

The first analysis in the field of musical ability was carried out by Drake who, when in London, had been a student of Spearman. He used a method of analysis developed by Spearman which tended to produce a general factor and specific factors, i.e. specific to each test. Wing, being a student of Burt's, used a method of factorisation developed by Burt which gives similar results to the centroid method used by Thurstone in America. The main difference between the method used by Wing and by McLeish, and those by American and Swedish psychologists who have tried to factorise musical ability tests is that the latter rotated their factors. The effect of this is to minimise the importance of general factors in favour of group factors.

The relative advantages and disadvantages of maximising the importance of a general factor over group or multiple factors have been discussed by Vernon (1950) particularly in relation to general intelligence. He points out that a group factor is almost infinitely subdivisible depending only on the degree of detail to which the analysis is carried. An important point is how broad a group factor should be before it can be accepted as a useful element in our picture of mental structure. The advantage of a broad general factor is that it is more stable as among different groups and depends less on the reliability of small groups of tests. Vernon, however, concludes that provided the form of factor analysis chosen accounts equally well for the original correlations by the means of the same limited number of factors, any one is equally legitimate.

The most extensive factorial study in the field of music so far reported was carried out by Holmstrom (1963). The Varimax method of rotation he adopted was intended to make the rotation of factors more objective.

Below are outlined the main factor loadings reported in the literature. Readers interested in inspecting the detailed tables for themselves will find most of them reproduced in Holmstrom's book both in their original form and after his re-analysis.

FIELDHOUSE 1937

Aim To investigate causes of singing out of tune.
Tests Seashore's; Mainwaring's Pitch and Rhythm; vocal range; auditory acuity; intelligence.
Subjects Fifty schoolchildren, aged nine to eleven, selected by their teachers as unable to sing in tune; with scores of five or less on eartests and who sang two songs of their own choosing off pitch.
Control group of 96 'normal' boys.

Unrotated factors

Special Group General factor: Seashore Pitch ·68; Intensity ·68; Mainwaring 'High-Low' ·65; Mainwaring Pitch ·59.
Normal Group General factor: Seashore Pitch ·73; Mainwaring Pitch ·68; Mainwaring 'High-Low' ·60; Auditory Acuity −·39.

Rotated factors:

Special Group:
I High-Low ·84; Mainwaring Pitch ·67; Seashore Pitch ·64; Memory ·57; vocal range ·40.
II Intensity ·93; Seashore Rhythm ·44.
III Mental Age ·42; Auditory Acuity −·48; vocal range −·49.

Normal Group:
I Mainwaring Pitch ·80; High-Low ·70; Seashore Pitch ·62; Memory ·60; vocal range ·42; Seashore Rhythm ·41; Intensity ·40.
II Seashore Rhythm ·40; Consonance ·40; Memory ·31.
III Mental Age ·52; Seashore Pitch ·47; Time ·43.
IV Mainwaring Rhythm ·50; Intensity ·41; Seashore Rhythm ·32.

DRAKE 1939

Tests used Drake Musical Memory and Retentivity; Seashore Pitch, Time, Rhythm, Intensity and Tonal Memory; K–D Tonal Movement.
Subjects 163 English boys, average age 13; unselected musically.
Factorial Technique. Spearman's Tetrad difference.

Unrotated factors

Common factor with over 30% of the variance, five group factors: tonal memory and tonal movement; pitch and intensity; pitch and tonal memory; musical memory and tonal memory; intensity and time.

Rotated factors

(Re-analysed by Karlin (1941) by centroid method and oblique rotation.)
X '*Memory*'. Rhythm ·531; Drake Memory ·510; Retentivity ·427.
Y '*Tonal sensitivity*'. Pitch ·572; Intensity ·513.
Z '*Retentivity*'. K–D Tonal Movement ·580; Seashore Memory ·542; Retentivity ·384.

Rotated factors

(As re-analysed by Holmstrom (1963) by J-method and Varimax rotation.)
I Intensity ·70; Pitch ·64; Drake Memory ·39.
II K–D Tonal Movement ·67; Seashore Memory ·66.
III Rhythm ·64; Drake Memory ·38; Seashore Memory ·30.
IV Retentivity ·59; Drake Memory ·52.

KARLIN 1941 STUDY

Tests used Drake Memory, Intervals and Retentivity; a simplified version of Seashore Pitch, Consonance, Rhythm, Time and Intensity; Emotional Sensitivity (which of eight adjectives best describes the emotional situation in a composition played).
Subjects 120 South African undergraduates.
Method Centroid.

Unrotated factors

Three factors (one general, two bipolar).
I Drake Memory ·732; Retentivity ·625; Intervals ·564; Pitch ·448; Rhythm ·443.
II Retentivity contrasted with Drake Memory.
III Retentivity and Rhythm v Interval Discrimination, Pitch and Time.

Rotated factors

I 'Memory for Form': Drake Memory ·561; Time ·414.
II 'Memory for Elements': Retentivity ·497; Rhythm ·398; Emotional Sensitivity ·318.
III 'Tonal Sensitivity': Intervals ·537; Time ·437; Pitch ·431.

Re-analysis by Holmstrom (1963)

By J-method and Varimax rotation
(Holmstrom included nine intellectual and literary tests which Karlin had correlated with the musical tests, but had not factorised because their correlations with the music variables were very low. Holmstrom reported eight factors. On the whole the musical variables had little in common with the intellectual and literary tests.)

I Intellectual and literary (Emotional Sensitivity ·30; Rhythm ·15).
II Drake Memory ·77; Retentivity ·58; Time ·55.
III Classification ·57; Intensity ·48.
IV Retentivity ·54; Intervals ·20.
V Synonyms ·46; Vocabulary ·31; Rhythm ·10.
VI Intervals ·58; Pitch ·54; Time ·47.
VII Rhythm ·36; Consonance ·32; Emotional Sensitivity ·32; Drake Memory ·29.
VIII Inferences ·48; Rhythm ·37; Analogies ·24.

KARLIN 1942 ANALYSIS

Tests used 31 auditory, auditory–verbal and visual tests, including Seashore 1939 battery and unfilled time (1919); intelligence.
Subjects 200 high-schoolchildren, aged 15 to 19.

Unrotated factors (Modified Centroid)

1 general factor (15%) Pitch (pure tones of short impulse) ·78; Pitch (vocal sounds) ·77; Seashore Pitch ·71; Memory ·58; Smallest Loading, Loudness (complex tones).

Rotated factors (Oblique Method)

A. 'Pitch–Quality' High loadings (·60 to ·70), all pitch except pitch discrimination of complex tones; Seashore Timbre ·44; and Memory ·42.
B. 'Loudness' All loudness and memory for male voices and some loading on time.
C. 'Auditory Integral' (i.e. loudness and time).
D. 'Auditory Resistance' (both to distortion and to masking effects).
E. 'Speed of Closure' (rapidity of perception).
F. 'Auditory Span Formation'. 'Actual mechanism by which auditory memory span is performed, having time, loudness and pitch elements'.
G. Memory (Seashore Memory ·36).
H. 'Incidental Closure', i.e. 'ability to survey extensive array of possible stimuli and attend to a selected few, so as to be able to reproduce crucial stimuli later'.
J. Uninterpretable.

Re-analysis by Holmstrom

Seven Seashore tests only.

Two factors

I Memory ·64; Pitch ·55; Rhythm ·47; Timbre ·47.
II Loudness ·60; Sense of Time (unfilled) ·58; Sense of Time (filled) ·47.

WING 1936

Tests used Chords, Rhythm, Phrasing (as in present battery); discords, note present or not in chords, notation reading, intervals, which notes move and melody dictation.
Subjects 33 boys, aged 11 to 13.
Method Centroid.

Unrotated factors

Five factors.

I General factor: Intervals ·779; Melody Dictation ·735; Chords ·645; Rhythm ·639.
II 'Power to analyse explicitly or implicitly the relations between musical stimuli': Chords ·473; Melody Dictation ·457; Notation −·542; Discords −·528.
III 'Power to retain auditory image': Notation Reading ·414; Discords ·394; Phrasing −·408; Chords −·372; Rhythm −·359.
IV 'Impression' or 'feeling' versus cognitive choice: Discords ·299; Melody Dictation ·219; Note in Chord −·319; Intervals −·282.
V 'Notational': Dictation ·166; Rhythm −·265.

In Holmstrom's re-analysis no fewer than seven factors were considered significant, but residual values were still too high to allow meaningful interpretation of the factors.

WING 1941

Main study

Tests used The seven Wing tests.
Subjects 43 boys aged 14 to 16.
Method Burt's weighted summation (re-analysis by Hotelling's method of principal components gave similar results).

Unrotated factors

(1 general 40·8%. 2 bipolar 13·4% and 3·1%).

I Highest loading: Phrasing ·765; Lowest, Rhythm ·421.

II Analytic v Synthetic. Chords, Pitch, Memory v Rhythm, Phrasing Intensity. Harmony zero.

III Harmony v Melody and Rhythm. Positive loadings: Harmony and Chords. Negative loadings: rest of battery.

Rotated factors

(Rotated to simple structure by Faulds)

I 'Melodic memory': Highest loading ·624; Memory, Pitch ·423; Rhythm and Harmony zero loadings.

II 'Qualitative judgments' (similar to Wing's Factor II).

III 'Harmony' (similar to Wing's Factor III).

Due to the relatively small number of subjects, no significant factors were found on Holmstrom's re-analysis.

Subsidiary study

Tests used The seven Wing tests and six others with which Wing was experimenting.
Subjects Not specified.
Method Burt's simple summation.

Unrotated factors

(One general, one bipolar)

I Intervals ·842, Rhythm ·842, Notation ·841; only test less than ·50 was intensity ·223.

II Intensity ·544; Discords ·380 v Pitch −·436 and Notation −·427.

VERNON 1950; brief report only

Tests used 17 tests including Oregon, Seashore Pitch, Rhythm, Memory and a Musical Knowledge test.
Subjects 70 students.

Unrotated factors

General Factor: Total Oregon ·84; Musical Knowledge ·84; Seashore Memory ·65; Seashore Pitch ·28; Seashore Rhythm ·35.
Group factor: 'The three Seashore tests had a considerable group factor of their own.'

WHITTINGTON 1957

Tests Wing Battery and Raven's Matrices (intelligence).
Subjects 24 musical New Zealand adolescents. 24 unmusical New Zealand adolescents.
Method Burt's summation.

Group I (musical)

Unrotated factors

1 general factor (44%).

I General factor: Rhythm ·81; Harmony ·77; Chords ·68; Memory ·67; Intensity ·63; Phrasing ·60; Pitch ·54; Intelligence ·58.

Group II (unmusical)

Unrotated factors

1 general factor (28%).

I General factor: Pitch ·71; Phrasing ·65; Intelligence ·62; Memory ·54; Rhythm ·52; Chords ·44; Intensity ·34; Harmony ·29.

MCLEISH 1950

Aim To test the validity of the Seashore tests against tests using musical material.
Tests Wing's battery; 1919 Seashore and the Oregon Discrimination.
Subjects 100 Psychology students.
Method Burt's simple summation.

Unrotated factors

Seashore (separately)

I General factor. Highest loading: Memory ·76; Pitch ·59; Rhythm ·27.

II Bipolar. Rhythm ·52 and Memory ·31 v 'the tests requiring immediate discrimination between sensory stimuli'.

Wing (separately)

I General (37%). Highest loading: Pitch ·76, Memory ·76; lowest Phrasing ·39.

II Bipolar (10%) 'interest–attention' factor Memory (·43); Harmony (·36) v Phrasing (−·49) and Rhythm (−·31).

III Bipolar (8%) 'Analytic–synthetic' Intensity (·48) v Rhythm (−·43).

Seashore, Wing and Oregon

I General (38·8%) Highest: Oregon 'Nature of change' ·86; Seashore Memory ·82; Wing Memory ·78; Wing Pitch ·78; Oregon 'preferred' ·77; Seashore Pitch ·66; lowest Seashore Time ·30.
II Bipolar (7%) Seashore ·33 and Wing ·30; Memory, Harmony ·28 v Seashore Pitch (−·49); Intensity (−·44); Wing Phrasing (−·39).

Holmstrom's re-analysis of Seashore analysis

Rotated factors

I Consonance ·69; Memory ·51.
II Rhythm ·71; Memory ·55.
III Intensity ·61; Pitch ·56; Time ·44.

FAULDS 1959

Aim To examine the sense of pitch in a variety of situations, some musical, some not, in order to investigate the relations between these areas.
Tests Seashore Pitch and Memory, Lundin Intervals, Wing Pitch, Franklin TMT group; 6 other auditory tests, e.g. octaves and scales played flat, sharp or correctly; Pitch and timbre discrimination with delay of 5–7 sec. between the two tones; auditory digit span.
Subjects 67 freshmen from Westminster Choir College (USA) (musical group) and 35 freshmen from Princeton (unselected musically).
Method Principal Axes; rotated by Extended Vectors.

Unrotated factors

General factor: Wing Pitch ·860; Seashore Memory ·842; Octaves ·778; auditory digit span ·015.

Rotated factors

I 'Music Factor' TMT ·560; Flat or sharp tunes ·518.
II 'Pitch': Lundin Intervals ·884; Wing Pitch ·700; Seashore Pitch ·655.
III 'Memory': Timbre (time delayed) ·788; Seashore Memory ·652; Pitch (time delayed) ·602.

FRANKLIN 1956

Study 1

Tests used Seashore Pitch and Memory; Wing Battery; Revesz–Franklin Rhythm; Franklin Individual TMT; Intelligence.

Subjects 79 women training-college students.
Method Centroid.

Unrotated factors

(One general, three bipolar).
General factor: Highest loading: Seashore Memory ·75; Wing Pitch ·71; Wing Memory ·68; Rev.–Fr. rhythm without music ·69; with music ·68; Intelligence ·14.

Rotated factors

(Successive Approximation method)
I Wing Memory ·59; Pitch ·54; Seashore Memory ·48; Pitch ·35.
II TMT ·68; Seashore Memory ·57; Wing Memory ·44; Harmony ·41; Pitch ·40; Rev.–Fr. rhythm with music ·35.
III Rev.–Fr. rhythm without music ·68; rhythm with music ·65; Intelligence ·53.
IV Loadings on all Wing tests except Memory.

Re-analysis by Holmstrom

(Varimax rotation)
I TMT ·75; Wing Harmony ·68; Seashore Memory ·65; Wing Pitch ·57; Wing Memory ·52; Rev.–Fr. rhythm with music ·43.
II Rev.–Fr. rhythm without music ·70; with music ·62; Intelligence ·55.
III All Wing except Memory.
IV Seashore Pitch ·64; Wing Memory ·53; Pitch ·48; Seashore Memory ·40.

Study 2

Tests Seashore Pitch and Memory; Wing Battery; Franklin TMT Individual and Group; Franklin Drummed and Melodic Rhythm; two tests of visual perception, one of intelligence and one of vocabulary (a second, retest score on Seashore Pitch included as a separate variable).
Subjects 157 training-college students.
Method Centroid.

Unrotated factors

General factor: Highest loading: Seashore Memory ·799; Wing Memory ·778; Seashore Pitch ·743 and ·728; Intensity ·251; Wing Rhythm ·244; Phrasing ·176; Intelligence ·404, ·406; Visual Perception ·141 and ·113.

Rotated factors

(by Simple Structure Method)

A. 'Pitch': Seashore Pitch ·58 and ·53; Wing Pitch ·25.
B. Visual Perception ·65 and ·40; Franklin Melodic Rhythm ·18; Wing Rhythm ·17.
C. Tonal Memory: Wing Memory ·32; Seashore Memory ·27.
D. Intelligence and vocabulary.
E. Franklin Drummed Rhythm ·50; Melodic Rhythm ·33.
F. Wing Intensity ·43.
G. Wing Memory ·52; Seashore Memory ·45; Chords ·41; Wing Pitch ·30.
H. Wing Phrasing ·43; Wing Memory ·21.
J. TMT group ·55; Intelligence ·20.

Holmstrom's re-analysis

a. with the retest Seashore Pitch, 16 significant factors.
b. without retest Seashore Pitch, 7 fully significant, one on verge of significance.

Both re-analyses produced substantially the same results as Franklin's original, except that factors F and H seemed to be combined into one factor which had a loading of ·45 on phrasing and ·41 on intensity, and the two tonal memory tests did not become a separate factor corresponding to Franklin's 'C' factor. Instead Wing's rhythm emerged as a largely specific factor.

HOLMSTROM 1963

Aim To investigate musical ability factors in a prognostic situation.
Tests Simplified version of Wing's tests 1–3; Holmstrom rhythm test; musical knowledge; attitude to school music; attitude to a programme of music; intelligence test; marks for reading, writing and arithmetic in grades 2 and 4. Criterion variable of school music marks in grades from 2 to 7.
Subjects 'E' study 184 children in Enköping schools.
'U' study 845 children in Uppsala schools (grades 6–7 analysis 325).
'UB' study 120 children selected from U group as coming from musically poor homes.
Method J-method with Varimax rotation.

Unrotated factors

not published

E study

Analysis 1. Marks for music in grade 2 compared with the other grade 2 variables. Analysis 2 marks for music grade 4 compared with the same variables. Analysis 3 marks for music in grade 4 compared with other grade 4 variables. Three very similar factors found in all 3 analyses.

E 1–2

I Intelligence and scholastic achievement. Moderate loading (·37 and ·39) on musical knowledge and on rhythm (·26).
II 'Musicality': Pitch ·59; Memory ·61 and ·58; Rhythm ·54 and ·53; Music Marks ·46; Chords ·32 or ·28; Intelligence ·27.
III 'Attitude to music': Musical Marks ·54; Attitude (school music) ·53; Attitude (programme) ·51.

E 3

I Intelligence and school marks.
II 'Music': Memory ·70; Pitch ·57; Rhythm ·55; Chords ·48; Intelligence only ·08.
III 'Attitude': Attitude (school music) ·55; Attitude (programme) ·39; Knowledge ·30; Music Marks ·24; Rhythm ·20.

U study

In four separate analyses, grade 2 variables compared with music marks in grades 2, 4, 6 and 7. In three more analyses, grade 4 variables were compared with music marks in grades 4, 6 and 7.

U 1–4

Five factors
I Intelligence, school subjects: musical knowledge (·39 to ·43); Rhythm (·12 to ·16); Pitch (·12 to ·15).
II 'Gamma' Music: High loadings on Rhythm and Memory; on Pitch (except U 4 ·23); Intelligence (·48; ·39 U 4).
III 'Alpha' Music: High loadings on Chords and Pitch; moderate on Memory; low on Rhythm. Loading on music marks increases from ·14 (U 1) to ·27 (U 3).
IV Attitude to music. High loadings on both attitude tests; increased loading on music marks (·30) in U 4.
V Marks for Music. High loadings on marks for music, with sizeable loadings (·43 to ·53 U 1 to U 3; ·32 U 4) on Memory.

U 5–7

Eight or ten factors not all interpreted by Holmstrom

I Intelligence and school subjects: musical knowledge (·23); Rhythm (·19 and ·21); Pitch (·18 and ·19).
II 'Beta' Music. High loadings on Memory and Pitch; moderate on Rhythm and Chords. High loadings on Music marks.
III 'Alpha' Music. High loadings on Chords (·61 to ·68); light loadings on Pitch and Rhythm and Music marks. Zero or ·1 on Memory.
IV High loadings only on programme; small loadings on Music marks and Rhythm.
V Marks for Music. Largely specific factor.

UB study

In three separate studies, grade 2 variables compared with music marks in grades 2, 4 and 6. In two analyses, grade 4 variables compared with music marks in grades 4 and 6.

UB 1–3

Three factors

I Intelligence – scholastic. *No* loading on musical knowledge. *Negative* loadings(– ·33 to –·37) on attitude to schoo lmusic.
II Attitude to music (both tests), marks for Music and Memory (light loadings on the other music tests).
III 'Gamma' Music. High loadings on all music tests, *except* Chords; high loadings on Musical Knowledge and Intelligence.

UB 4–5

Five factors

I Intelligence – scholastic.
II 'Beta' Music. High loadings on Memory, Pitch and Chords; moderate on Rhythm. Music marks ·46.
III Attitude (both) and marks for Music and Memory ·33.
IV Rhythm ·50; Intelligence ·43; Arithmetic ·31. (U 5, U 4 similar.)
V Knowledge of Music ·64; Memory ·34; Rhythm ·26. (U 5, U 4 similar).

SHUTER 1964

Aim To compare the results obtained on the Wing battery with five different groups. Groups 1 and 2 were similar in age and general educational background, but differed in expertness in handling music. Group 3 were similar to group 1 in being highly musical, but differed in age. Groups 4

and 5 were comparable in all respects, including level of musical ability, but differed in that group 4 were all men and group 5 entirely women.
Tests The Wing battery.

Subjects

Group 1. 41 students of the Eastman School of Music, Rochester, USA.
Group 2. 100 training college students, 48 men and 52 women, whose scores were close to the adult average on Wing's norms.
Group 3. 100 members of the National Youth Orchestra (aged 12 to 18).
Group 4. 100 men training-college students with Wing scores ranging from 78 to 110, i.e. of above average ability.
Group 5. 100 women training-college students of similar musical level to group 4.
Method Principal Components.

GROUP 1

Unrotated factors

(General Factor 34·59%; 4 bipolar factors, 15·55%, 14·55%, 11·67%, 9·58%)
I Chords ·753; Memory ·569; Pitch ·470; Intensity ·466; Rhythm ·449; Phrasing ·388; Harmony ·296.
II Phrasing ·690; Intensity ·635 v Memory −·533.
III Harmony ·803; Rhythm ·520 v Memory −·434.
IV Pitch ·793; Harmony ·260 v Rhythm −·418.
V Rhythm ·480; Phrasing ·358 v Chords −·407.

Rotated factors

(The three most significant factors rotated)
I Memory ·878; Chords ·693; Pitch ·478; Rhythm ·246; Harmony, Intensity and Phrasing zero.
II Phrasing ·790; Intensity ·787; Chords ·319.
III Harmony ·838; Rhythm ·672; Chords ·139.

GROUP 2

Unrotated factors

(1 broad factor 25·56%, 4 bipolars 17·58%, 13·75%, 12·26% and 11·28%)
I Memory ·598; Pitch ·511; Harmony ·500; Phrasing ·364; Chords ·320; Rhythm and Phrasing zero.
II Intensity ·701; Phrasing ·485; Rhythm ·354; Harmony ·258 v Pitch −·609; Memory −·262.

III Chords ·692; Harmony ·174 v Rhythm −·688; Pitch −·231.
IV Chords ·552; Rhythm ·419 v Memory −·530; Intensity −·415.
V Harmony ·641; Rhythm ·322 v Phrasing −·549.

GROUP 3

Unrotated factors

(1 broad factor 32·28%, 4 bipolars 16·16%, 13·26%, 13·04%, 11·27%)

I Memory ·778; Chords ·728; Harmony ·637; Rhythm ·564; Pitch ·543; Intensity ·324; Phrasing −·013.
(As the children were probably tired after a long day's rehearsing, when tested by Wing, it seemed likely that fatigue had affected their performance of the last test. Shuter therefore believed that factor I should be considered a general factor, the zero loading for phrasing being due to the test situation.)
II Phrasing ·879; Rhythm ·422 v Intensity −·394.
III Intensity ·571; Pitch ·544 v Chords −·454.
IV Intensity ·609; Harmony ·363 v Pitch −·495.
V Rhythm ·619 v Harmony −·503.

GROUP 4

Unrotated factors

(1 broad factor 35·49%, 4 bipolars 15·47%, 13·35%, 12·43%, 8·99%)

I Highest loading: Harmony ·749; Pitch ·684; Chords ·521; Phrasing ·496; Memory ·490; Rhythm ·314; Intensity −·017.
II Rhythm ·653; Intensity ·512; Phrasing ·469 v Memory −·444.
III Intensity ·798 v Rhythm −·401.
IV Chords ·709 v Memory −·608.
V Phrasing ·585 v Rhythm −·493.

GROUP 5

Unrotated factors

(1 broad factor 35·28%, 5 bipolars 31·81%, 12·65%, 11·87%, 10·05%, 8·56%)

I Highest loading: Pitch ·653; all other loadings ·5 except Phrasing ·343 and Rhythm ·021.
II Rhythm ·767; Harmony ·403; Chords ·280 v Phrasing −·394.
III Rhythm ·545; Memory ·498 v Harmony −·428 and Chords −·405.
IV Phrasing ·737 v Memory −·439 and Pitch −·359.
V Intensity ·661 v Harmony −·446 and Phrasing −·313.
VI Pitch ·589 v Memory −·480.

APPENDIX III

Correlations between Intelligence Tests and Musical Ability Tests

(In almost every case the intelligence test has been a group test)

Abbreviations: P Pitch; I Intensity; T Time; C Consonance; M Memory; R Rhythm; Ca Cadence; Ph Phrasing; Ch Chords; H Harmony; PR Pitch Recognition; MR Melody Recognition.

SEASHORE MEASURES

Investigator	Subjects	P	I	T	C	M	R
Weaver (1924)	94 college students	·35	·24	·12	·06	·26	—
Fracker & Howard (1928)	230 college students	·32	·01	·13	·09	·10	·12
Highsmith (1929)	59 female music school students	·58	·35	·39	−·14	·30	—
Salisbury & Smith (1929)	131 training college students	·31	·15	·30	·00	·24	·02
	144 training college students	·39	—	·49	·38	·33	·24
Farnsworth (1931)	150 University students	·14	·11	·10	−·38	·11	·17
Drake (1940)	163 boys; age = 13	·12	·14	·08	·03	·07	·05
Franklin (1956)	(*a*) 79 training college students	·13				·00	
	(*b*) 157 training college students	·15				·14	
Hollingworth (1926)	49 children with IQs above 135	Median PR 46·7	50·0	58·0		52·3	
Manor (1950)	4th grade children	·21				·27	·11[1]
Christy (1956)	103 college students	·18				·18	·33[2]

[1] 1939 Revision, Form A. [2] 1939 Revision, Form B.

KWALWASSER–DYKEMA TESTS

Investigator	Subjects	Full battery
Newkirk[3]	1,000	·34
Robertson[3]	Over 5,000 children aged 8 to 20	·33
Lambert[3]	1,024 children aged 11	·33
Lehman (1950)	450 musicians & college students	·18
Chase[3]	82 feeble-minded children (IQ range 45–77)	Average PR = 35·0 (tests 1–8)
Drake (1940)	As above	·06 (Melodic Taste) ·13 (Tonal Movement)

KWALWASSER MUSIC TEST

Radley[3]	550 children	·51
Bentley, R. (1955)	(*a*) 87 instrument-playing music students	·34
	(*b*) 95 non-instrument-playing music students	·46

[3] Cited by Kwalwasser, 1955.

LOWERY TESTS

Investigator	Subjects	M	Ca	Ph
Lowery (1929)	Group of school girls aged 12–14	·44	·44	·00
Drake (1940)	As above		·06	

MAINWARING TESTS

Investigator	Subjects	P	R	M
Mainwaring (1931)	83 Elementary schoolchildren	·53	·46	·04
	34 grammar school boys	·39	·32	—

DRAKE TESTS

Investigator	Subjects	M	R
Drake (1957)	158 college students	·28	
Drake	163 music students (aged 7–16)	·27	
Drake	20 high school children	·05	·10
Drake	61 psych. students		·00
Drake	130 students		—·03
Drake	130 students		·05
Karlin (1941)	120 students	·06	
Christy (1956)	103 students	·21	

OREGON MUSIC DISCRIMINATION TEST

Investigator	Subjects	
Hevner (1931)	74 college students	—·15 (2 version form)
	74 college students	—·17 (4 version form)

WING TESTS OF MUSICAL INTELLIGENCE

Investigator	Subjects	Full Battery	Ch	P	M	R	H	I	Ph
Wing (1948)	23 girls	·30							
	42 boys	·32							
	24 adults	·40							
	24 adults	·34							
	454 college students	·20							
Edmunds (1960)	Sec. Mod. school-children:								
	60 A & D stream		—·07	·33	·39		·28 (tests 1–3)		
	58 F stream & ESN		—·02	·36	·47		·39 (tests 1–3)		
Coulthard (1952)	32 grammar school boys	·04							
Shuter (1964)	200 Royal Marine School of Music boys		·180 (tests 1–3)			·154 (tests 4–7)			
Whittington (1957)	24 musical adolescents		·36	·18	·42	·40	·47	·52	·20
	24 unmusical adolescents		·21	·63	·32	·20	·17	·00	·40
Bentley, R.	As above (*a*)	·39	·21	·39	·37	·22	·18	·01	·02
	(*b*)	·39	·22	·39	·26	·03	·25	·17	·11
Franklin	As above (*a*)		·09	—·10	—·02	·00	—·19	·20	·04
	(*b*)		·09	·12	·20	·23	·21	—·03	·08
Parker (1961)	1,174 high school children								
	Boys				·134 (tests 4–7)				
	Girls				·054				
	Both sexes				·075				

WING–HOLMSTROM TESTS

Investigator	Subjects	Ch	P	M	R
Holmstrom (1963)	189 children in Grade 2	·25	·28	·23	·33
	765 children in Grade 2	·16	·32	·29	·33
	651 children in Grade 4	·09	·17	·22	·27
	120 children from unmusical homes				
	Grade 2	·00	·17	·20	·34
	Grade 4	·14	·32	·16	·47

LUNDIN TESTS

Investigator	Subjects		Non-language	Language
Lundin (1949)	113 music students	·15	·25	·13
	155 unselected students	·24	·22	·19

WHISTLER–THORPE TESTS

Investigator	Subjects	R	PR	MR	P	R
Bentley	As above (*a*)	·25	·22	·32	·35	·26
	(*b*)	·00	·24	·22	·32	·01

GASTON TEST

Investigator	Subjects	Tests 19–23	Tests 24–28	Tests 29–33	Tests 34–40
Bentley	As above (*a*)	·15	·32	·29	·29
	(*b*)	·11	·18	·16	·25

FRANKLIN TMT TESTS

Investigator	Subjects	Group	Individual
Franklin	As above (*a*)	—	−·11
	(*b*)	·18	·18

BENTLEY MEASURES OF MUSICAL ABILITY

Investigator	Subjects	P	M	Ch	R
Bentley, A. (1966)	166 children aged 10 to 12	·30	·25	·24	·34
	149 children with IQs of 100 or above; age = 11:1	assoc. signif. at 1% level	no significant association		

APPENDIX IV

Correlations between Tests of Musical and other Abilities

Investigator	Subjects	Music Tests	Other Tests	*r* or range of correlation
Miscellaneous				
Franklin (1956)	157 TC students	Seashore P & M Wing 1–7;	Vocabulary	−·01 to ·28
		Franklin TMT & Rhythm	Visual Perception	−·11 to ·23
Coulthard (1952)	32 high school boys	Wing 1–7	Oral French	·53
		Wing 1–3	Oral French	·42
Edmunds (1960)	(*a*) 60 Sec. Mod. A & D stream	Wing 1–3	Reading	·33
	(*b*) 58 Sec. Mod. F stream & ESN			·24
Drake (1940)	24 to 186 women college students	Drake Memory	Various college subjects	−·13 to ·24
Holmstrom (1963)	(*a*) About 1,000 children aged 8 to 10	Wing–Holmstrom	Reading marks	·06 to ·37
	(*b*) 120 from un-musical homes	Wing–Holmstrom	Reading marks	−·02 to ·28

	(*a*)	Wing–Holmstrom	Writing marks	·09 to ·36
	(*b*)	Wing–Holmstrom	Writing marks	—·02 to ·27
Shuter (1964)	200 RMSM junior musicians	Wing 1–3	Spelling	·12
		Wing 4–7	Spelling	·12
			Spatial	·21
			Spatial	—·03
Other arts				
Carroll (1932)	133 college students	Oregon	Prose Appreciation	·12
				·13[1]
			Art Judgment	·16
				·15[1]
Rigg (1937)	71 male students	Oregon	Poetry	·34
				·24[1]
Morrow (1938)	112 male psychology students	K–D	Art	·10 average
Williams, Winter & Wood (1938)	Over 200 children aged 11 to 17	Appreciation	Literary	·26
				·18[1]
Karlin (1941)	120 college students	Drake Memory, etc. (see page 297)	Poetry Appreciation	·12
Drake (1957)	19 music students	Drake Rhythm	Art	·00
	166 Belgian boys			·14
Mathematic/Scientific				
Morrow (1941)	80 college students	Seashore	Arithmetic	—·09 to ·19
			Thurstone number series	—·08 to ·41
Edmunds (1960)	As (*a*) above	Wing 1–3	Arithmetic	·11
	As (*b*) above			·29
Holmstrom (1963)	As (*a*) above	Wing–Holmstrom	Arithmetic marks	—·01 to ·29
	As (*b*) above			—·09 to ·26
Shuter (1964)	200 RMSM junior musicians	Wing 1–3	Mathematics	·07
		Wing 4–7		·05

[1] With intelligence held constant.

BIBLIOGRAPHY OF REFERENCES CITED

AGNEW, M. (1922) 'The auditory images of great composers', *Psychol. Monogr.* **31**, 279–287.

ALLPORT, G. W. (1937) *Personality: A Psychological Interpretation.* New York: Holt.

ALVIN, J. (1965) *Music for the Handicapped Child.* London: OUP.

AMARIA, R., BIRAN, L. A. and LEITH, G. O. M. (1966) Learning in pairs. Research notes on programmed learning, No. 10. National Centre for Programmed Learning.

ANASTASI, A. and LEVEE, R. (1959) 'Intellectual defect and musical talent: a case report', *Amer. J. ment. Defic.*, **64**, 695–703.

ANDREWS, J. A. and WARDIAN, J. F. (1964) *Introduction to Music Fundamentals.* New York: Appleton-Century-Crofts.

ANTRIM, D. K. (1945) 'Do musical talents have higher intelligence?', *Etude*, **63**, 127–128.

ASHFORD, T. H. A. (1966) 'The use of programed instruction to teach fundamental concepts in music theory', *J. Res. Mus. Ed.*, **14**, 171–177.

ASHMAN, R. (1952) 'The inheritance of simple musical memory', *J. Hered.*, **43**, 51–52.

BACHEM, A. (1940) 'The genesis of absolute pitch', *J. Acoust. Soc. Amer.*, **11**, 434–439.

BAHLE, J. (1934) 'Die Gestaltübertragung im vokalen Schaffen zeitgenössischer Komponisten', *Arch. ges. Psychol.*, **91**, 444–451.

BARNES, R. A. (1964) *Fundamentals of Music: A Program for Self-Instruction.* New York: McGraw-Hill.

BAUMANN, V. H. (1960) 'Teen-age music preferences', *J. Res. Mus. Ed.*, **8**, 75–84.

BEKESY, G. VON (1957) 'The ear', *Sci. Amer.*, **197**, Aug., 66–78.

BELAIEW-EXEMPLARSRY, S. (1926) 'Das musikalische Empfinden im Verschulalter', *Zsch f. Ang. Psychol.*, **27**, 177–216.

BENTLEY, A. (1954) 'Children who sing out of tune', *Music in Education*, March/April, 13–14.

BENTLEY, A. (1957) 'Children who sing out of tune', *Music in Education*, July/Aug., 82–83, Sept./Oct., 114.

BENTLEY, A. (1966) *Musical Ability in Children and Its Measurement.* London: Harrap.

BENTLEY, R. R. (1955) A critical comparison of certain aspects of musical aptitude tests. Ph.D. thesis, Univ. Southern California.

BERGAN, H. A. (1957) A study of drop-outs in instrumental music in five selected schools in Michigan. Ed.D. thesis, Michigan State University.

BERGAN, J. R. (1966) 'Factors affecting pitch discrimination', *Bull. Council Res. Mus. Ed.*, No. 8, 15–21.

BIENSTOCK, S. F. (1942) 'A predictive study of musical achievement', *J. Genet. Psychol.*, **61**, 135–145.

BIGGS, M. R. (1960) An evaluation of technique employing the use of the magnetic tape recorder in teaching students of brass instruments. Ph.D. thesis, State Univ. of Iowa.

BILLROTH, T. (1895) *Wer ist musikalish?* Berlin: Ed. Hansick.

BLACKETER-SIMMONDS, D. A. (1953) 'An investigation into the supposed differences existing between mongols and other mental defectives with regard to certain psychological traits', *J. Ment. Sci.*, **99**, 702–719.

BLAND, D., ed. (1936) *Songs and pictures of a child.* London: Williams and Morgan.

BOTEZ, M. I. and WERTHEIM, N. (1959) 'Expressive aphasia and amusia following right frontal lesion in a right-handed man', *Brain*, **82**, 186–202.

BREHMER, F. (1925) 'Melodieauffassung und melodische Begabung in Kindes', *Z. angew. Psychol. Beihefte*, **36.**

BRENNAN, F. (1926) 'The relation between musical capacity and performance', *Psychol. Monogr.*, 36:1 (Whole No. 167), 190–248.

BRIDGER, W. H. (1961) 'Sensory habituation and discrimination in the human neonate', *Amer. J. Psychiat.*, **117**, 991–996.

BROADBENT, D. E. (1958) *Perception and communication.* London: Pergamon.

BRODY, V. (1953) 'The role of body-awareness in the emergence of musical ability: its application to music education, the college basic music course, and critic teaching', *J. Res. Mus. Ed.*, **1**, 21–29.

BROWN, R. W. (1936) 'The relation between age (chronological and mental) and rate of piano learning', *J. Appl. Psychol.*, **20**, 511–516.

BUGG, E. and HERPEL, L. (1946) 'The significance of tonal memory for musicality', *J. General Psychol.*, **35**, 3–15.

BÜHLER, C. (1935) *From Birth to Maturity: An Outline of the Psychological Development of the Child.* Kegan, Paul, Trench and Trubner.

BURKS, B. S., JENSEN, D. W. and TERMAN, L. M. (1930) *The Promise of Youth: Follow-up studies of a thousand gifted children.* Vol. III *Genetic studies of Genius.* Stanford, Calif.: Stanford Univ. Press.

BUROS, O. K. (1949) *The Third Mental Measurements Yearbook.* Highland Park, N.J.: Gryphon Press.

BUROS, O. K. (1953) *The Fourth Mental Measurements Yearbook.* Highland Park, N.J.: Gryphon Press.

BUROS, O. K. (1959) *The Fifth Mental Measurements Yearbook.* Highland Park, N.J.: Gryphon Press.

BUROS, O. K. (1966) *The Sixth Mental Measurements Yearbook*. Highland Park, N.J.: Gryphon Press.

BURROUGHS, G. E. R. and MORRIS, J. N. (1962) 'Factors involved in learning a simple musical theme', *Brit. J. Educ. Psychol.*, **32**, (1), 18–28.

BURT, C. (1909) 'Experimental tests of general intelligence', *Brit. J. Psychol.*, **3**, 94–177.

BURT, C. (1943) 'Ability and income', *Brit. J. Educ. Psychol.*, **13**, 83–98.

BURT, C. *et al.* (1945) 'The psychology of art' in *How the Mind Works*, 2nd edn. London: Allen & Unwin.

BURT, C. (1958) *The Backward Child*, 4th edn. London: ULP.

CAIN, M. L. (1960) A comparison of the Wing Standardised Tests of Musical Intelligence with a Test of Musicality by Gaston and the Drake Musical Aptitude Tests. Master's thesis, Univ. Kansas.

CANTOR, G. N. and GIRARDEAU, F. L. (1959) 'Rhythmic discrimination ability in mongoloid children', *Amer. J. ment. Defic.*, **63**, 621–625.

CAPURSO, A. A. (1934) 'The effect of an associative technique in teaching pitch and interval discrimination', *J. Appl. Psychol.*, **18**, 811–818.

CARLSEN, J. C. (1964) 'Programed learning in melodic dictation', *J. Res. Mus. Ed.*, **12**, 139–148.

CARLSEN, J. C. (1965) *Melodic Perception: A Program for Self-Instruction*. New York: McGraw Hill.

CARLSEN, J. C. (1965*a*) 'The role of programmed instruction in the development of musical skills'. Paper to Comprehensive Musicianship Seminar, Foundation for College Education in Music, Northwestern University.

CARROLL, H. A. (1932) 'A preliminary report on a study of the interrelations of certain appreciations', *J. Educ. Psychol.*, **23**, 505–510.

CARTER, C. O. (1962) *Human Heredity*. Harmondsworth: Penguin.

CASEY, G. J. (1964) A study of instrumental music drop-outs in the Moline (Illinois) schools. Ed.D. thesis, Music, Colorado State College.

CATTELL, R. B. (1950) *Personality: a systematic, theoretical and factual study*. New York: McGraw Hill.

CHERRY, C. (1953) 'Some experiments on the recognition of speech in one and two ears', *J. acoust. Soc. Amer.*, **25**, 975.

CHRISMAN, M. L. (1962) Preparation and experience of the music teacher in higher education. Ed.D. thesis, Education, Univ. Southern California.

CHRIST, W. B. (1953) The reading of rhythm notation approached experimentally according to the techniques and principles of word reading. Doctor's thesis, Indiana University.

CHRISTY, L. J. (1956) A study of the relationships between musicality, intelligence, and achievement. Ph.D. thesis, Music, Indiana University.

CLEAK, R. E. (1958) A study of musical aptitudes of children in a secondary school. Dip.Ed. thesis, Bristol University.

CLOUGH, J. (1964) *Scales, Intervals, Keys and Triads: A Self-Instruction Program*. New York: Norton.

COCHRAN, M. (1930) 'Kinesthesis and the piano', *Austral. J. Psychol.*, **8**, 205–209.

COFFMAN, A. R. (1951) The effect of training on rhythm discrimination and rhythmic action. Doctor's thesis, Northwestern University.

COLBY, M. G. (1935) 'Instrumental reproduction of melody by preschool children', *J. Genet. Psychol.*, **47**, 413–430.

COLWELL, R. (1963) 'An investigation of musical achievement, among vocal students, vocal-instrumental students and instrumental students', *J. Res. Mus. Ed.*, **11**, 123–130.

COLWELL, R. J. and RUNDELL, G. (1965) 'An evaluation of achievement in auditory–visual discrimination resulting from specific types of musical experiences among junior high school students', *J. Res. Mus. Ed.*, **13**, 239–245.

CONNETTE, E. (1941) 'The effect of practice with knowledge of results', *J. Educ. Psychol.*, **32**, 523–532.

COOPER, M. (1965) 'The value of practice', *Daily Telegraph and Morning Post*, 6 Feb. 1965, p. 11.

CORTOT, A. (1935) 'Do infant prodigies become great musicians?' *Music and Letters*, **16**, 2, 124–128.

COULTHARD, J. W. (1952) An investigation into certain aspects of oral ability in French and musical ability as measured by the Wing Test of Musical Intelligence. Ed.B. thesis, Aberdeen University.

COWELL, H. (1926) 'The process of musical creation', *Amer. J. Psychol.*, **37**, 233–236.

COX, C. (1926) *Genetic Studies of Genius:* Vol. II *The Early Mental Traits of Three Hundred Geniuses*. Stanford, Calif.: Stanford Univ. Press.

CRAMER, W. (1958) The relation of maturation and other factors to achievement in beginning instrumental music performance at the fourth through eighth grade levels. Ed.D. thesis, Music, Florida State University.

CROWDER, N. A. (1960) 'Automatic tutoring by intrinsic programming', pp. 286–298 in *Teaching Machines and Programmed Learning*, ed. A. A. Lumsdaine and R. Glaser. Washington: National Education Association.

CULPEPPER, L. R. (1961) A study of the hearing impairments in defective singers. Ed.D. thesis, Music, George Peabody College for Teachers.

CULVER, F. (1965) A study of the musical aptitude profile. M.A. thesis, Iowa University.

DALY, D. D. and BARRY, M. J. (1957) 'Musicogenic epilepsy', *Psychosom. Med.*, **19**, 399–408.

DAVENPORT, C. B. and STEGGERDA, M. (1929) *Race Crossing in Jamaica*. Washington: Carnegie Institution.

DELACATO, C. H. (1959) *The Treatment and Prevention of Reading Problems*. Illinois: Thomas.

DOIG, D. (1941–2) 'Creative music', *J. Educ. Res.*, **35**, 263–275; **35**, 344–355; **36**, 241–255.

DRAKE, R. M. (1931) Tests of musical talent. Ph.D. thesis, London University.

DRAKE, R. M. (1933) 'Four new tests of musical talent', *J. Appl. Psychol.*, **17**, 136–147; 'The validity and reliability of tests of musical talent', *J. Appl. Psychol.*, **17**, 447–458.

DRAKE, R. M. (1939) 'Factorial analysis of music tests by the Spearman-Tetrad-Difference Technique', *J. Musicol.*, **1**, No. 1, 6–16.

DRAKE, R. M. (1940) 'The relation of musical talent to intelligence and success at school', *J. Musicol.*, **2**, No. 1.

DRAKE, R. M. (1945) 'The effect of ear training on musical talent scores', *J. Musicol.*, **4**, No. 3, 110–112.

DRAKE, R. M. (1957) *Manual for the Drake Musical Aptitude Tests*. Chicago: Science Research Associates.

DREXLER, E. N. (1938) 'A study of the development of the ability to carry a melody at the preschool level', *Child Devel.*, **9**, 319–332.

DRINKWATER, H. (1916) 'Inheritance of artistic and musical ability', *J. Genet.*, **5**, 229–241.

DURKIN, H. (1937) 'Trial-and-error, gradual analysis, and sudden reorganisation', *Arch. Psychol.*, **30**, No. 210.

EASTMAN SCHOOL OF MUSIC (1966) Press release 19 June 1966, 'Suzuki coming to Eastman'.

ECCLES, J. C. (1958) 'The physiology of the imagination', *Sci. Amer.*, **199**, Sept., 135–146.

EDMUNDS, C. B. (1960) Musical ability, intelligence and attainment of Secondary Modern and E.S.N. children', Thesis, Leeds University.

EELLS, W. (1933) 'Mechanical, physical and musical ability of the native races of Alaska', *J. Appl. Psychol.*, **17**, 493–506.

FARNSWORTH, P. R. (1928) 'The effects of nature and nurture on musicality'. In the *27th Yearbook of the National Society for Education*, Part II, 233–247.

FARNSWORTH, P. R. (1931) 'An historical, critical and experimental study of the Seashore–Kwalwasser Test Battery', *Genet. Psychol. Monogr.*, **9**: 5, 291–393.

FARNSWORTH, P. R. (1938) 'Auditory acuity and musical ability in the first four grades', *J. Psychol.*, **6**, 95–98.

FARNSWORTH, P. R. (1941) 'Further data on the Adlerian theory of artistry', *J. Gen. Psychol.*, **24**, 447–450.

FARNSWORTH, P. R. (1958) *The Social Psychology of Music*. New York: Dryden Press.

FARNSWORTH, P. R. (1961) 'The use and abuse of music talent tests', *Amer. String Teacher*, **11** (2), 5–6.

FARNSWORTH, P. R. (1964) 'New weights for the Seashore–Hevner tests for attitude toward music', *Rev. Psychol. Music*, **3**, 1–6.

FAULDS, B. (1959) The perception of pitch in music. Princeton University, Princeton, N.J.

FEIS, O. (1910) *Studien über die Genealogie und Psychologie der Musikers.* Wiesbaden: Bergmann.

FERRELL, J. W. (1961) A validity investigation of the Drake Musical Aptitude Tests. Ph.D. thesis, Music, State Univ. of Iowa.

FIELDHOUSE, A. E. (1937) A study of backwardness in singing among school children. Ph.D. thesis, London University.

FOOT, E. (1965) The effect of the intensity of auditory feedback on the loudness–intensity, quality wave form, and intelligibility of the singer's voice. Ph.D. thesis, Music, Univ. Kansas.

FORBES, H. S. and FORBES, H. B. (1927) 'Fetal sense reaction: Hearing', *J. Comp. Psychol.*, **7**, 353–355.

FOSHA, L. (1964) A study of the validity of the Musical Aptitude Profile. Ph.D. thesis, State Univ. of Iowa.

FRACKER, G. C. and HOWARD, V. M. (1928) 'Correlation between intelligence and musical talent', *Psychol. Monogr.*, **39**, 2, No. 178, 157–161.

FRANKLIN, E. (1956) *Tonality as a Basis for the Study of Musical Talent.* Göteberg: Gumperts Forlag.

FRED, B. G. (1956) The instructional values of an exploratory string instrument in a fourth grade. Doctoral dissertation, Northwestern University.

FRIEND, R. (1939) 'Influences of heredity and musical environment on the scores of kindergarten children on the Seashore Measures of Musical Ability', *J. Appl. Psychol.*, **23**, 347–357.

FRY, D. (1948) 'An experimental study of tone deafness', *Speech*, **12** (1), 4–11.

GALTON, F. (1869) *Hereditary Genius.* London: Macmillan.

GASTON, E. T. (1958) *A test of musicality, manual of directions.* Lawrence, Kansas: Odell's Instrumental Service.

GEDDA, L. *et al.* (1961) 'L'eredità della attitudini musicali', Proc. 2nd Internat. Congr. of Human Genetics, Rome.

GESELL, A. and ILG, F. (1943) *The infant and child in the culture of today.* London: Hamilton.

GESELL, A. and ILG, F. (1946) *The child from five to ten.* London: Hamilton.

GILBERT, G. M. (1942) 'Sex differences in musical aptitude and training', *J. Gen. Psychol.*, **26**, 19–33.

GILBERT, J. A. (1893) 'Experiments on the musical sensitiveness of school-children', *Studies from the Yale Psychological Laboratory* 1892–3, **1**, 80–87.

GORDON, E. (1961) 'A study to determine the effects of training and practice on Drake Musical Aptitude Test scores', *J. Res. in Mus. Ed.*, **9** (1), 63–74.

GORDON, E. (1965) *Musical Aptitude Profile Manual*. Boston: Houghton Mifflin.

GORDON, K. (1917) 'Some tests on the memorising of musical themes', *J. Exp. Psychol.*, **2**, 93–99.

GOULD, A. O. (1965) 'The development of specialised programs for singing in the elementary school', *Bul. Council for Res. in Mus. Ed.*, No. 6, 17–24.

GRANT, W. (1963) *Music in education*. Proc. 14th symposium of the Colston Research Society. London: Butterworth.

GRAVES, W. L. (1963) A comparison of three methods for improving intonation in the performance of instrumental music. Ed.D. thesis, Univ. Colorado.

GRAVES, W. S. (1947) 'Factors associated with children's taking music lessons, including some parent–child relationships', *J. Genet. Psychol.*, **7**, 65–89 and 91–125.

GRAVES, W. S. (1949) Adolescents' musical training as a function of parent/child relationships and associated variables. Doctoral dissertation, Columbia University.

GROVE, G. (1954) *Dictionary of Music and Musicians*, 5th edn. ed. Eric Blom. Macmillan.

GUILFORD, J. P. (1957) 'Creative abilities in the arts', *Psychol. Rev.*, **64**, 110–118.

HALLER, M. W. (1932) 'The reactions to changes in the intensity and pitch of pure tones', *J. Genet. Psychol.*, **40**, 162–180.

HAMMER, H. (1963) 'An experimental study of the use of the teaching of melodic sight singing', *J. Res. Mus. Ed.*, **11**, 44–54.

HARDER, P. (1965) *Basic Materials in Music Theory: A Programed Course*. Boston: Allyn and Bacon.

HATTWICK, M. S. and WILLIAMS, H. M. (1935) 'The measurement of musical development II', *Univ. Iowa Stud. Child Welfare*, **11**, 1–100.

HEARNSHAW, L. S. (1951) 'Exploring the intellect', *Brit. J. Psychol.*, **62**, 315–321.

HEBB, D. O. (1949) *The organisation of behavior*. New York: Wiley.

HEIM, K. (1963) Musical aptitude of senior high school students in residential schools for the blind. M.M. thesis, Education, Univ. Kansas.

HEINLEIN, C. P. (1929) 'A new method of studying the rhythmic responses of children', *J. Genet. Psychol.*, **36**, 205–228.

HELLER, J. J. (1962) The effects of formal music training on the Wing Musical Intelligence scores. Ph.D. thesis, Music, State Univ. of Iowa.

HENDERSON, M. T. (1931) Remedial measures in motor rhythm as applied to piano performance. Thesis, Iowa University.

HENKIN, R. I. (1955) 'A factorial study of the components of music', *J. Psychol.*, **39**, 161–181.

HEVNER, K. (1931) 'A study of tests for appreciation of music', *J. Appl. Psychol.*, **15**, 575–583.

HEVNER, K. (1937) 'The aesthetic experience: a psychological description', *Psychol. Rev.*, **44**, 245–263.

HIGHSMITH, J. A. (1929) 'Selecting musical talent', *J. Appl. Psychol.*, **13**, 486–493.

HILLBRAND, E. K. (1923) *Hillbrand Sight-Singing Test.* Yonkers, N.Y.: World Book.

HIRIARTBORDE, E. (1964) *Les Aptitudes Rythmiques.* Paris: Dactylo-Sorbonne.

HOFFREN, J. (1964) 'The construction and validation of a test of expressive phrasing in music', *J. Res. Mus. Ed.*, **12**, 159–164.

HOLLINGWORTH, L. S. (1926) 'The musical sensitivity of children who test above 135 IQ', *J. Ed. Psych.* 17, 95–109.

HOLMES, J. (1954) 'Increased reliabilities, new keys and norms for a modified Kwalwasser–Dykema test of musical aptitudes', *J. Genet. Psychol.*, **85**, 65–73.

HOLMES, J. (1954*a*) 'A substrata analysis of spelling ability for elements of auditory images', *J. Exp. Educ.*, **22**, 329–349.

HOLMSTROM, L. G. (1963) *Musicality and Prognosis.* Uppsala: Almqvist and Wiksells.

HORACEK, L. (1963) 'Programmed instruction in music', *Rev. Psychol. Music*, No. 2.

HORBULEWICZ, J. (1963) The development of musical memory. Doctoral thesis, Higher School of Education, Danzig.

HOWES, F. (1958) *Music and its Meanings.* Three Lectures delivered in the University of London. London: Athlone.

HUDSON, L. (1966) *Contrary Imaginations.* London: Methuen.

HURST, C. C. (1912) 'Mendelian heredity in man', *Eugen. Rev.*, **4** (1), 20–24.

HUXLEY, A. L. (1960) *The Doors of Perception.* London: Chatto & Windus.

IHRKE, W. R. (1963) 'Automated music training', *J. Res. Mus. Ed.*, **11** (1), 3–10.

IHRKE, W. R. (1966) 'Automated rhythm training (progress report)', *Bul. Council Res. Mus. Ed.*, No. 7, 34–37.

IL'INA, G. A. (1959) 'On the formation of musical conceptions in preschool children', *Vop. Psikhol.*, **5** (5), 134–144.

IL'INA, G. A. (1961) 'Characteristics of the development of musical rhythm in children', *Vop. Psikhol.*, No. 1, 119–132.

JAMIESON, R. P. G. (1951) An investigation into songs known by Scottish schoolchildren and their musical preferences. Ed.B. thesis, Glasgow University. Summary in *Brit. J. Educ. Psychol.*, **22** (1), 74–75.

JERSILD, A. T. and BIENSTOCK, S. F. (1931) 'The influence of training on the vocal ability of three-year-old children', *Child Devel.*, **2**, 272–290.
JERSILD, A. T. and BIENSTOCK, S. F. (1934) 'A study of the development of children's ability to sing', *J. Educ. Psychol.*, **25**, 481–503.
JERSILD, A. T. and BIENSTOCK, S. F. (1935) 'The development of rhythm in young children', *Child Devel. Monogr.*, No. 22, Teachers' College, Columbia.
KALMUS, H. (1949) 'Tune deafness and its inheritance', *Proc. Int. Cong. Genetics*, Stockholm, 605.
KALMUS, H. (1952) 'Inherited sense defects', *Sci. Amer.*, **186** (5), 64–70.
KANABLE, B. M. (1965) An experimental study comparing programmed instruction with classroom teaching of sight singing. Ph.D. thesis, Music, Northwestern University.
KARLIN, J. E. (1941) 'Music ability', *Psychometrika*, **6**, 61–65.
KARLIN, J. E. (1942) 'A factorial study of auditory function', *Psychometrika*, **7**, 251–279.
KIMURA, D. (1964) 'Left–right differences in the perception of melodies', *Quart. J. Exp. Psychol.*, **16**, 355–358.
KIRKPATRICK, W. C. (1962) Relationships between the singing ability of pre-kindergarten children and their home musical environment. Ed.D. thesis, Education, Univ. Southern California.
KLEIST, K. (1962) *Sensory aphasia and amusia; the myeloarchitectonic basis*. New York: Pergamon Press.
KLINEBERG, O. (1935) *Race differences*. New York: Harper.
KOTICK, M. L. and TORGERSON, T. L. (1950) *Diagnostic tests of achievement in music*. Los Angeles: Calif. Test Bureau.
KRIES, J. VON (1926) *Wer ist musikalisch?* Berlin.
KUDER, G. F. (1956) *Kuder Preference Record: Vocational*. Chicago: Science Research Associates.
KUHN, W. E. and ALLVIN, R. (1967) 'Computer-assisted teaching: a new approach to research in music', *Bul. Council Res. Mus. Educ.*, No. 11, 1–13.
KWALWASSER, J. (1955) *Exploring the Musical Mind*. New York: Coleman-Ross.
KWALWASSER, J. and RUCH, G. (1924) *Kwalwasser–Ruch Tests of Musical Accomplishment*. Iowa City: Bureau of Ed. Res. and Serv.
KYME, G. H. (1956) 'Are musical tastes indicative of musical capacity?', *J. Res. Mus. Ed.*, **4**, 44–51.
LAMBERT, C. (1934) *Music Ho! A study of music in decline*. London: Faber & Faber.
LARSON, R. (1955) 'Finding and guiding musical talent', *Music Educ. J.*, **42**, Sept., 22–25.
LASHLEY, K. (1950) 'In search of the engram', *Symposia of the Society for Experimental Biology*, IV, 454–482.

LAWTON, A. (1933) *Foundations of Practical Ear Training.* Oxford University Press. London: Humphrey Milford.

LEHMAN, C. F. (1950) 'A comparative study of instrumental musicians on the basis of the Kwalwasser–Dykema Music Tests, the Otis Intelligence Test and the Minnesota Multiphasic Personality Inventory', *J. Educ. Res.*, **44**, 57–61.

LEITH, G. O. M. and BURKE, K. M. (1966) Mode of responding and redundancy. Nat. Centre for Programmed Learning Research Notes No. 7.

LEONTIEV, A. N. (1957) 'The nature and formation of human psychic properties', pp. 226–232 in *Psychology in the Soviet Union* (ed. B. Simon). London: Routledge and Kegan Paul.

LONG, N. H. (1965) A revision of the University of Oregon Music Discrimination Test. Ph.D. thesis, Indiana University.

LOWERY, H. (1926) 'Cadence and phrase tests in music', *Brit. J. Psychol.*, **17**, 111–118.

LOWERY, H. (1929) 'Musical memory', *Brit. J. Psychol.*, **19**, 397–404.

LOWERY, H. (1952) *The Background of Music.* London: Hutchinson's University Library.

LUNDIN, R. W. (1944) 'A preliminary report on some new musical ability tests', *J. appl. Psychol.*, **28**, 393–396.

LUNDIN, R. W. (1949) 'The development and validation of a set of musical ability tests', *Psychol. Monogr.*, **63**: 305, 1–20.

LUNDIN, R. W. (1958) 'What next in the psychology of musical measurement?' *Psychol. Rec.*, **8**, 1–6.

LUNDIN, R. W. (1963) 'Can perfect pitch be learned?', *Music educ. J.*, **49**, 49–51.

LUNDIN, R. W. (1967) *An objective psychology of music.* New York: Ronald.

MCGEOCH, J. A. (1942) *The psychology of human learning.* New York: Longmans, Green.

MCGINNIS, E. M. (1928) 'Seashore's measures of musical ability applied to children of the pre-school age', *Amer. J. Psychol.*, **40**, 620–623.

MCLEISH, J. (1950) 'The validation of Seashore's measures of musical talent by factorial methods', *Brit. J. Psychol. (Stat. Section)*, **3**, 129–140.

MADISON, T. H. (1942) 'Interval discrimination as a measure of musical aptitude', *Arch. Psychol.*, **268**, p. 100.

MAINWARING, J. (1931) 'Experiments on the analysis of cognitive processes in musical ability', *Brit. J. Educ. Psychol.*, **1**, 180–203; 'Tests of musical ability', *Brit. J. Educ. Psychol.*, **1**, 313–321.

MAINWARING, J. (1933) 'Kinaesthetic factors in the recall of musical experience', *Brit. J. Psychol.*, **23**, 284–307.

MAINWARING, J. (1947) 'The assessment of musical ability', *Brit. J. Educ. Psychol.*, **17**, 83–96.

MAINWARING, J. (1948) 'Gestalt psychology applied to music' in Grove's *Dictionary of Music and Musicians.*

MALTZMANN, E. (1964) An investigation of key-tone matching with children and adults. Ed.D. thesis, Music, Boston Univ. School of Educ.

MALTZMANN, E. (1965) 'Research on "Analysis of musical behavior" ', *Programed Instruction*, **4**, No. 9, 1, 4, 5 and 12.

MANN, C. W. (1940) 'Mental measurements in primitive communities', *Psychol. Bul.*, **37**, 366–395.

MANOR, H. C. (1950) 'A study in prognosis', *J. Educ. Psychol.*, **41**, 31–50.

MARQUIS, J. H. (1963) A study of interval problems in sight-singing performance with consideration of the effect of context. Ph.D. thesis, Music, State Univ. of Iowa.

MARTIGNETTI, A. J. (1965) 'Causes of elementary instrumental music drop-outs', *J. Res. Mus. Ed.*, **13**, 177–183.

MARTIN, G. (1966) *Basic concepts in music*, Belmont, Calif.: Wadeworth.

MILLER, G. (1956) 'The magical number seven plus or minus two', *Psychol. Rev.*, **63**, 81–97.

MILLER, G. A. (1964) *Psychology, the science of mental life*. London: Hutchinson.

MILLER, G. A. and SELFRIDGE, J. A. (1950) 'Verbal content and recall of meaningful material', *Amer. J. Psychol.*, **63**, 176–185.

MILNER, B. (1962) 'Laterality effects in audition'. In Mountcastle, V. (ed.), *Interhemispheric relations and cerebral dominance*. Baltimore.

MINISTRY OF EDUCATION (1960) *Teaching Music in the Schools*. Pamphlet No. 27, HMSO.

MINOGUE, B. (1923) 'A case of secondary mental deficiency with musical talent', *J. Appl. Psychol.*, **7**, 349–352.

MJOEN, J. (1926) 'Genius as a Biological Problem', *Eugen. Rev.*, **17**, 242–257.

MJOEN, J. (1934) *Die Vererbung der musikalischen Begabung*. Berlin: Alfred Metzner.

MONROE, W. S. (1903) 'Tone perception and music interest of young children', *Ped. Sem.*, **10**, 144–146.

MOORHEAD, G. E. and POND, D. (1941) *The Music of Young Children: I Chant*. Pillsbury Found. Stud.

MOORHEAD, G. E. and POND, D. (1942) *The Music of Young Children: II General Observations*. Pillsbury Found. Stud.

MOORHEAD, G. E., SANDVIK, F. and WIGHT, D. (1951) *The Music of Young Children: IV Free Use of Instruments for Musical Growth*. Pillsbury Found. Stud.

MORGAN, C. T. (1965) *Physiological Psychology*. New York: McGraw-Hill.

MORROW, R. S. (1938) 'An analysis of the relations among tests of musical, artistic and mechanical abilities', *J. Psychol.*, **5**, 253–263.

MORROW, R. S. (1941) 'An experimental analysis of the theory of independent abilities', *J. educ. Psychol.*, **32**, 494–512.

MOSHER, R. M. (1925) *Mosher Test of Individual Singing*, Teach. Coll.

Contrib. Educ., No. 194. N.Y. Bureau of Publications, Teachers Coll., Columbia University.

MUELLER, J. H. *et al.* (1934) 'Studies in appreciation of art', 4, No. 6, 1–151.

MUELLER, KATE HEVNER, (1956) 'Studies in music appreciation', *J. Res. Mus. Ed.*, 4, 3–25.

MULL, H. K. (1925) 'The acquisition of absolute pitch', *Amer. J. Psychol.*, **36**, 469–493.

MURSELL, J. L. (1937) *The Psychology of Music.* New York: Norton.

MURSELL, J. L. (1948) *Education for Musical Growth.* Boston: Ginn & Co.

MURSELL, J. L. and GLENN, M. (1931) *The Psychology of School Music Teaching.* New York: Silver Burdett.

MYERS, C. S. (1922) 'Individual differences in listening to music', *Brit. J. Psychol.*, **13**, 52–71.

NEFF, W. D. (1964) 'Temporal pattern discrimination in lower animals and its relation to language perception'. In *Disorders of Language*, Ciba Foundation Symposium, ed. de Reuck and O'Connor.

NELSON, C. B. (1956) 'Follow-up of an experimental teaching method in music at the fourth and fifth grade levels', *J. Exp. Educ.*, **24**, 283–289.

NELSON, C. B. (1967) 'The effectiveness of the use of adjunct programed analyses of musical works on students' perception of form', *Bul. Council Res. in Mus. Educ.*, **9**, 29–43.

NEWTON, G. DE C. (1959) *Selection of Junior Musicians for Royal Marines School of Music: An Evaluation of H. D. Wing's Test.* Senior Psychologist's Department, Admiralty.

NIELSON, J. T. (1930) 'A study in the Seashore Motor Rhythm test', *Psychol. Monogr.*, **40**, 74–84.

NORTHRUP, W. C. (1931) 'The inheritance of musical ability', Student Pedigree Studies, *Eugen. News*, **16.**

OLIVER, R. (1932) 'The musical talent of natives in East Africa', *Brit. J. Psychol.*, **22**, 333–343.

OWENS, W. A. and GRIMM, W. (1941) 'A note regarding exceptional musical ability in a low-grade imbecile', *J. Educ. Psychol.*, **32**, 636–637.

PARKER, O. G. (1961) A study of the relationship of aesthetic sensitivity to musical ability, intelligence, and socioeconomic status. Ed.D. thesis, Music, Univ. Kansas.

PARTHASARATHY, K. (1957) 'Musical aptitude and appreciation among high school students', *J. Psychol. Res.* (Mysore), **1** (2), 49–59.

PELLETIER, H. W. (1963) An investigation of the relation between training in instrumental music and selected aspects of language growth in third grade children. Ed.D. thesis, Education, Arizona State University.

PENFIELD, W. (1958) 'Functional localization in temporal and deep sylvian areas'. In *The Brain and Human Behavior.* Baltimore: Williams and Wilkins. pp. 210–226.

PETZOLD, R. G. (1960) 'The perception of music symbols in music reading by normal children and by children gifted musically', *J. Exp. Educ.*, **28**, 271–319.

PETZOLD, R. G. (1966) *Auditory perception of musical sounds by children in the first six grades.* Cooperative Research Project No. 1051. Univ. of Wisconsin.

PFLEDERER, M. (1964) 'The responses of children to musical tasks embodying Piaget's principles of conservation', *J. Res. Mus. Ed.*, **12**, 251–268.

PFLEDERER, M. (1966) 'How children conceptually organize sounds', *Bul. Council Res. in Mus. Educ.*, No. 7, 1–12.

PFLEDERER, M. and SECHREST, L. (1967) *How children conceptually organize musical sounds.* Cooperative Research Project No. 6-10-285. Northwestern University.

PITMAN, D. J. (1965) 'The musical ability of blind children', *Rev. Psychol. in Music*, **2** (2), 19–28.

PLATT, W. (1933) 'Temperament and disposition revealed in young children's music', *Character and Personality*, **2**, 246–251.

POLLOCK, T. (1950) Singing disability in school children. M.Ed. thesis, Durham University.

POSKANZER, D. C., BROWN, A. E. and MILLER, H. (1962) 'Musiogenic epilepsy caused only by a discrete frequency band of church bells', *Brain*, **85**, 77–92.

PRATT, C. C. (1928) 'Quarter-tone music', *J. genet. Psychol.*, **35**, 286–293.

PREYER, W. (1901) *The Mind of the Child.* Part I: *The Senses and the Will.* New York: Appleton.

PRONKO, N. H. and BOWLES, J. W. (1951) *Empirical foundations of psychology.* New York: Holt, Rinehart and Winston.

RAINBOW, E. L. (1965) 'A pilot study to investigate the constructs of musical aptitude', *J. Res. Mus. Ed.*, **13**, 3–14.

RAVEN, J. C. (1957) *A guide to using Progressive Matrices Sets A, Ab, and B.* London: Lewis.

RESER, H. (1935) 'Inheritance of musical ability', Student Pedigree Studies, *Eugen. News*, **20**.

REVESZ, G. (1920) 'Prüfung der Musikalität', *Zsch. f. Psychol.*, **85**, 163–209.

REVESZ, G. (1925) *The Psychology of a Musical Prodigy.* New York: Harcourt Brace.

REVESZ, G. (1946) 'Beziehung zwischen mathematischer und musikalischer Begabung', *Schweiz Zsch. f. Psychol.*, **5**, 269–281.

REVESZ, G. (1953) *Introduction to the Psychology of Music.* London: Longmans, Green.

RICHET, G. (1900) 'Note sur un cas remarquable de précocité musicale', IV Congress internationale de Psychologie.

RIESEN, A. H. (1950) 'Arrested vision', *Sci. Amer.*, 183, July, 16–19.

RIFE, D. C. and SNYDER, L. H. (1931) 'A genetic refutation of the principles of behavioristic psychology', *Hum. Biol.*, **3**, 547–559.

RIGG, M. (1937) 'The relationship between discrimination in music and discrimination in poetry', *J. Educ. Psychol.*, **28**, 149–152.

RITCHIE, T. V. (1960) A study of the effects of diatonic harmony upon the aural perception of selected music fragments. Ph.D. thesis, Music, Indiana University.

ROBY A. R. (1962) 'A study in the correlation of music theory grades with the Seashore Measures of Musical Talents and the Aliferis Music Achievement Test', *J. Res. Mus. Ed.*, **10**, 137–142.

RODERICK, J. L. (1965) An investigation of selected factors of the creative thinking ability of music majors in a teacher training program. Ed.D. thesis, Music, Univ. Illinois.

ROGERS, V. R. (1956) Children's expressed musical preferences at selected grade levels. Ed.D. thesis, Music, Syracuse University.

ROSS, F. B. (1914) 'The measurement of time sense as an element in the sense of rhythm', *Psychol., Monogr.*, **16**, 166–172.

ROWE, A. W. (1959) *The Education of the Average Child.* London: Harrap.

RUPP, H. 'Über die Prüfung musikalischer Fähigkeiten', (Teil I), *Zsch. ang. Psychol.*, **9**, 1–76.

SACHS, C. (1957) *The Lore of Non-Western Music.* Essay in *Some Aspects of Musicology.* New York: Liberal Arts Press.

SAKURABAYASHI, H., SATO, Y. and UEHARA, E. (1956) 'Auditory discrimination of the blind', *J. Psychol., Blind*, **1**, 3–10.

SALISBURY, F. S. and SMITH, H. R. (1929) 'Prognosis of sight-singing ability of normal school students', *J. Appl. Psychol.*, **13**, 425–439.

SANDERSON, H. E. (1933) 'Differences in musical ability in children of different national and racial origin', *J. Genet. Psychol.*, **42**, 100–120.

SCHEERER, M., ROTHMANN, E. and GOLDSTEIN, K. (1945) 'A case of "Idiot Savant": an experimental study of personality organization', *Psychol. Monogr.*, **58** (4), pp. 64.

SCHEINFELD, A. (1956) *The New Heredity and You.* London: Chatto & Windus.

SCHOEN, M. (1923) 'The validity of tests of musical talent', *J. Comp. Psychol.*, **3**, 101–121.

SCHOEN, M. (1925) 'Tests of musical feeling and musical understanding', *J. comp. Psychol.*, **5**, 31–52.

SCHOEN, M. (1940) *The Psychology of Music.* New York: Roland Press.

SEARS, M. (1965) 'The tape recorder employed in the development of children's singing: an experimental study', *Colorado J. Res. Mus. Educ.*, **11**, 8–12.

SEASHORE, C. E. (1919) *The Psychology of Musical Talent.* New York: Silver Burdett.

SEASHORE, C. E. (1937) 'The psychology of music: XI', *Mus. Ed. J.*, **24** (3), 25–26.

SEASHORE, C. E. (1938) *Psychology of Music*. New York: McGraw-Hill.

SEASHORE, C. E., LEWIS, D. and SAETVEIT, J. C. (1960) *Manual of Instructions and Interpretations for the Seashore Measures of Musical Talents*, 2nd Revision. New York: The Psychological Corporation.

SEASHORE, C. E. and LING, T. L. (1918) 'The comparative sensitiveness of blind and seeing persons', *Psychol. Monogr.*, **25**, 148–158.

SEASHORE, C. E. and MOUNT, G. H. (1918) 'Correlation of factors in musical talent and training', *Psychol. Monogr.*, **25**, 47–92.

SEASHORE, R. (1926) 'Studies in motor rhythm', *Iowa Stud., Psychol.*, **9**, 142–199.

SEASHORE, R. H. (1935) 'Improvability of pitch discrimination', *Psychol. Bull.*, **32**, 546.

SEMEONOFF, B. (1940) 'Further developments in a new approach to the testing of musical ability, with special reference to groups of secondary school children', *Brit. J. Psychol.*, **31**, 145–161.

SENDON, M. VON. (1960) *Space and Sight: The Perception of Space and Shape in the Congenitally Blind before and after operation*. Trans. Heath. London: Methuen.

SERGEANT, D. (1967) 'The incidence and characteristics of absolute pitch amongst musicians'. Paper delivered to the 3rd Conf. on Res. in Mus. Educ., Reading University.

SHIELDS, J. (1962) *Monozygotic Twins Brought up Apart and Brought up Together*. London: OUP.

SHINN, M. W. (1907) *The Development of the Senses in the First Three Years of Childhood*. Univ. Calif. Publ. in Educ., Vol. 4.

SHIRLEY, M. M. (1933) *The First Two Years: II Intellectual Development*. Univ. Minnesota Press.

SHUTER, R. P. G. (1964) An investigation of hereditary and environmental factors in musical ability. Ph.D. thesis, London. University.

SHUTER, R. P. G. (1966) 'Hereditary and environmental factors in musical ability', *The Eugen. Rev.*, **58**, 149–156.

SIEVERS, D. J. (1955) Development and standardization of a test of psycholinguistic growth in preschool children. Ph.D. thesis, Univ. Illinois.

SIMONS, G. M. (1964) 'Comparisons of incipent music responses among very young twins and singletons', *J. Res. Mus. Ed.*, **12**, 212–226.

SKINNER, B. F. (1961) 'Teaching machines', *Sci. Amer.*, **205**, Nov., 90–102.

SKINNER, B. F. (1963) 'B. F. Skinner in Oxford', *Bull. Brit. Psychol. Soc.*, **16**, No. 50, 9–11. Reprint of Report, printed in *Nature*, on the 'Future of Teaching Machines'.

SLONIMSKY, N. (1948) 'Musical children, prodigies or monsters?', *Etude*, **66**, 591–592.

SMITH, F. O. (1914) 'The effect of training in pitch discrimination', *Psychol. Monogr.*, **17** (3), Whole No. 69, 67–103.

SMITH, O. (1957) 'Relationship of rhythm discrimination to motor rhythm performance', *J. appl. Psychol.*, **41**, 365–369.

SMITH, R. A. (1965) 'Patterns of meaning in aesthetic education: with special reference to the teaching of music and art at the secondary level (7–12)', *Bul. Council Res. Mus. Educ.*, No. 5, 1–12.

SMITH, R. B. (1963) 'The effects of group vocal training on the singing ability of nursery school children', *J. Res. Mus. Educ.*, **11**, 137–141.

SPOHN, C. L. (1959) An exploration in the use of recorded teaching material to develop aural comprehension in college music classes. Ph.D. thesis, Education, Ohio State University.

SPOHN, C. L. (1963) 'Programming the basic materials for self-instructional development of aural skills', *J. Res. Mus. Ed.*, **11**, 91–98.

SPOHN, C. L. (1965) A comparison between different stimuli combined with two methods of providing knowledge of results in music education. Ohio State Univ. Research Foundation.

SPOHN, C. L. and POLAND, W. (1964) *An evaluation of two methods using magnetic tape recordings of programed instruction in the elemental materials of music.* Ohio State Univ. Research Foundation. Final Report RF.1407.

STAFFORD, R. E. (1965) Nonparametric analysis of twin data with the Mann–Whitney U test. Research Report No. 10, Louisville Twin Study Child Development Unit, Univ. Louisville School of Medicine.

STAMBACK, M. (1960) *Trois epreuves de rythme.* In *Manuel pour l'examen psychologique de l'enfant*, ed. R. Zazzo. Neuchatel: Editions Delachaux et Niestle.

STANDING CONFERENCE FOR AMATEUR MUSIC (1966) *Music and the Newsom Report: A Theme with eight variations*, Ref. No. 717. London: Nat. Council of Social Services.

STANTON, H. (1922) 'Inheritance of specific musical capacities', *Psychol. Monogr.*, **31** (1), Whole No. 140, 157–204.

STANTON, H. (1935) 'Measurement of musical talent: the Eastman Experiment', *Univ. Iowa Stud. Psychol. Music*, **2**, 1–140.

STANTON, H. and KOERTH, W. (1930) 'Musical capacity measures in adults repeated after musical education', *Univ. Iowa Stud. Aims Prog. Res.*, **31**.

STANTON, H. and KOERTH, W. (1935) 'Musical capacity measures in children repeated after musical training', *Univ. Iowa Stud. Aims Prog. Res.*, **42**: 259.

STENDHAL (HENRI BEYLE) (1956) *Life of Rossini*, trans. Coe. London: Calder.

STERN, W. (1924) *The psychology of early childhood up to the sixth year of age*, trans. Barwell. New York: Henry Holt.

STOKES, C. F. (1944) An experimental study of tachistoscopic training in reading music. Doctoral thesis, Univ. Cincinnati.

STRONG, E. K. (1959) *Vocational Interest Blank for Men, Vocational Interest Blank for Women.* Stanford, Calif.: Stanford Univ. Press.

STUMPF, C. (1883–90) *Tonpsychologie.* Leipzig: Hirzel.

STUMPF, C. (1909) 'Akustische Versuche mit Pepito Ariola', *Z. ang. Psychol.*, **2**, 1–11.

SWARD, K. (1933) 'Jewish musicality in America', *J. Appl. Psychol.*, **15**, 675–712.

TARRELL, V. V. (1965) 'An investigation of the validity of the musical aptitude profile', *J. Res. Mus. Ed.*, **13**, 195–206.

TAYLOR, E. M. (1941) 'A study of the prognosis of musical talent', *J. Exp. Psychol.*, **10**, 1–28.

TAYLOR, I. A. (1959) 'The nature of the creative process', In Smith, P. (ed.), *Creativity: an examination of the creative process.* New York: Hastings House.

TEPLOV, B. M. and BORISSOVA, M. N. (1957) 'La sensibilité différentielle et la memoire sensorielle', *Questions de psychologie*, **1**.

TEPLOV, B. M. (1966) *Psychologie des Aptitudes Musicales*, trans. Deprun. Paris: Presses Universitaires de France.

TERMAN, L. M. *et al.* (1925) *Genetic Studies of Genius:* Vol. I. *The Mental and Physical Traits of a Thousand Gifted Children.* Stanford, Calif.: Stanford Univ. Press.

THACKRAY, R. (1966) Rhythmic abilities. Paper to 2nd Conf. on Research in Music Education.

THORPE, W. H. (1956) 'The language of birds', *Sci. Amer.*, **195**, Oct., 129–138.

TILLIS, M. (1960) *Chords and Discords: The Life of an Orchestral Musician.* London: Phoenix House.

TOMATIS, A. (1953) 'L'Oreille Musicale', *J. Francais d'Oto-Rhino-Laryngologie*, No. 2.

TREDGOLD, A. F. (1922) *Mental Deficiency (Amentia)*, 4th ed. London: Bailliere, Tindall and Cox.

TREDGOLD, R. and SODDY, K. (1956) *A Textbook of Mental Deficiency*, 9th edn. London: Bailliere, Tindall and Cox.

TRIESMANN, A. (1966) 'Our limited attention', *Adv. of science*, **22**, 600–611.

TYSON, M. (1966) 'Creativity'. In *New Horizons in Psychology*, ed. Foss, B. Harmondsworth: Penguin.

UPDEGRAFF, R., HEILEGER, L. and LEARNED, J. (1938) 'The effect of training upon the singing ability and musical interest of three-, four-, and five-year-old children', *Univ. of Iowa Stud. Child Welfare*, **14**, 83–121.

VALENTINE, C. W. (1956) *The Normal Child.* Harmondsworth: Penguin Books.

VALENTINE, C. W. (1962) *The Experimental Psychology of Beauty.* London: Methuen.

VAN ALSTYNE, D. and OSBORNE, E. (1937) 'Rhythm responses of Negro and White children two to six', *Monogr. Soc. Res. Child Devel.*, **2**, 4.

VANDENBERG, S. G. (1962) 'The hereditary abilities study: hereditary components in a psychological test battery', *Am. J. Hum. Genet.*, **14** (2), 220–237.

VATER, H. (1934) 'Musikalische Produktion', *Arch. ges. Psychol.*, **90**, 1–60.

VERNON, M. D. (1960) 'The development of perception in children', *Educ. Res.*, **3** (1), 2.

VERNON, P. E. (1931) The psychology of music with especial reference to its appreciation, perception and composition. Ph.D. dissertation, Cambridge University.

VERNON, P. E. (1933) 'The apprehension and cognition of music', *Proc. Royal Music Assoc.*, **59**, 61–84.

VERNON, P. E. (1934–5) 'Auditory Perception: I The Gestalt Approach', *Brit. J. Psychol.*, **25** (2), 123–139. 'Auditory Perception: II The Evolutionary Approach', *Brit. J. Psychol.*, **25** (3), 265–281.

VERNON, P. E. (1950) *The Structure of Human Abilities.* London: Methuen.

VERNON, P. E. (1960) *Intelligence and attainment tests.* London: Univ. London Press.

VIDOR, M. (1931) *Was ist Musikalitat?* Munich: Beck.

WALLACE, W. (1914) *The Musical Faculty.* London: Macmillan.

WALTER, BRUNO (1961) *Of Music and Music-making*, trans. Hamburger. London: Faber and Faber.

WALTER, W. GREY (1953) *The living brain.* London: Duckworth.

WARDIAN, J. F. (1963) An experiment concerning the effectiveness of programmed learning for use in teaching the fundamentals of music. Ed.D. thesis, Education, Washington State University.

WATKINS, J. G. and FARNUM, S. E. (1954) *The Watkins–Farnum Performance Scale.* Winona, Minn.: Leonard Music.

WATSON, J. B. (1930) *Behaviorism*, rev. edn. New York: Norton.

WEAVER, A. T. (1924) 'Experimental studies in vocal expression: II The prediction of talent for vocal expression in reading', *J. Appl. Psychol.*, **8**, 159–186.

WEINERT, L. (1929) 'Untersuchungen über das absolute Gehör', *Arch. ges. Psychol.*, **73**, 1–128.

WERTHEIM, N. and BOTEZ, M. I. (1961) 'Receptive amusia', *Brain*, **84**, 19–30.

WHIPPLE, G. M. (1901) 'An analytical study of the memory image and the process of judgments in the discrimination of clangs and tones', *Amer. J. Psychol.*, **12**, 409–457.

WHIPPLE, G. M. (1903) 'Studies in pitch discrimination', *Amer. J. Psychol.*, **14**, 289–309.

WHISTLER, H. S. and THORPE, L. P. (1950) *Musical Aptitude Test.* Los Angeles: Calif. Test Bureau.

WHITE, A. (1961) 'The Aliferis Music Achievement test as a predictor of success in music theory at St Olaf College', *J. Ed. Res.*, **54**, 315–317.

WHITE, R. K. (1931) 'The versatility of genius', *J. Soc. Psychol.*, **2**, 460–489.

WHITTAKER, W. G., HUTCHISON, W. O. and PICKFORD, R. W. (1942) 'Symposium on the psychology of music and painting', *Brit. J. Psychol.*, **33**, 40–57.

WHITTINGTON, R. W. T. (1957) 'The assessment of potential musical ability in secondary school children', *J. Educ. Psychol.*, **48**, 1–10.

WHYBREW, W. (1962) *Measurement and evaluation in music.* Dubuque, Iowa: Brown.

WILLIAMS, A. A. (1958) 'Number readiness', *Ed. Rev. Inst. Educ. Birmingham*, **11**, 31–46.

WILLIAMS, E. D., WINTER, L. and WOOD, J. M. (1938) 'Tests of literary appreciation', *Brit. J. Ed. Psychol.*, **8**, 265–284.

WILLIAMS, N. M., SIEVERS, C. H. and HATTWICK, M. S. (1933) 'The Measurement of Musical Development', *Univ. Iowa Stud. in Child Welfare*, VII, No. 1, pp. 191.

WILSON, M. E. (1951) *How to help your child with music.* New York: Schuman.

WING, H. D. (1936) Tests of musical ability in school children. M.A. thesis, London University.

WING, H. D. (1941) 'A factorial study of musical tests', *Brit. J. Psychol.*, **31**, 341–355.

WING, H. D. (1941*a*) Musical ability and appreciation. Ph.D. thesis, London University.

WING, H. D. (1948) 'Tests of musical ability and appreciation', *Brit. J. Psychol. Monogr. Suppl.*, No. 27, pp. 88.

WING, H. D. (1954) 'Some applications of test results to education in music', *Brit. J. Educ. Psychol.*, **24**, 161–170.

WING, H. D. (1955) 'Musical aptitude and intelligence', *Education to-day*, **5**, No. 1.

WING, H. D. (1960) *Manual for Standardised Tests of Musical Intelligence* (City of Sheffield Training College, Ref. 600), N.F. Ed. Res.

WING, H. D. (1962) 'A revision of the Wing musical aptitude test', *J. Res. Mus. Ed.*, **10** (1), 39–46.

WING, H. D. (1963) 'Is musical aptitude innate?', *Rev. Psychol. in Music*, **1**, 1–7.

WITHERSON, M. I. (1935) The effect of race and nationality on music measurement. Master's thesis, Syracuse University.

WOLFF, P. H. (1963) 'Observations on the early development of smiling', *Determinants of Infant Behaviour:* II, ed. Foss. London: Methuen.

WOLNER, M. and PYLE, W. H. (1933) 'An experiment in individual training in pitch-deficient children', *J. Educ. Psychol.*, **24**, 602–608.

WRIGHT, F. A. (1928) 'The correlation between achievement and capacity in music', *J. Educ. Res.*, **17**, 50–56.

WYATT, R. F. (1945) 'The improvability of pitch discrimination', *Psychol. Monogr.*, **58**: 267, 1–58.

YATES, N. and BRASH, H. (1941) 'An investigation of physical and mental characteristics of a pair of twins reared apart from infancy', *Ann. Eugen.*, **11**, 89–101.

ZELIG, T. (1967) 'A direct approach to pre-school violin teaching', *Mus. Ed. J.*, **53**, Jan., 44–46.

Author Index

General Index

(including titles of tests)