Psychology *of* Music

From Sound to Significance

SECOND EDITION

Siu-Lan Tan, Peter Pfordresher and Rom Harré

ROUTLEDGE

Psychology of Music

In *Psychology of Music: From Sound to Significance*, Second Edition, the authors consider music on a broad scale, from its beginning as an acoustical signal to its different manifestations across cultures. In their second edition, the authors apply the same richness of depth and scope that was a hallmark of the first edition of this text. In addition, having laid out the topography of the field in the original book, the second edition puts greater emphasis on linking academic learning to real-world contexts, and on including compelling topics that appeal to students' natural curiosity. Chapters have been updated with approximately 500 new citations to reflect advances in the field.

The organization of the book remains the same as the first edition, while chapters have been updated and often expanded with new topics. 'Part I: Foundations' explores the acoustics of sound, the auditory system, and responses to music in the brain. 'Part II: The Perception and Cognition of Music' focuses on how we process pitch, melody, meter, rhythm, and musical structure. 'Part III: Development, Learning, and Performance' describes how musical capacities and skills unfold, beginning before birth and extending to the advanced and expert musician. And finally, 'Part IV: The Meaning and Significance of Music' explores social, emotional, philosophical, and cultural dimensions of music and meaning.

This book will be invaluable to undergraduate and postgraduate students in psychology and music, and will appeal to anyone who is interested in the vital and expanding field of psychology of music.

Siu-Lan Tan is Professor of Psychology at Kalamazoo College in Michigan, USA. She completed degrees in Music at Pacific Union College, graduate studies at Oxford University, and a PhD in Psychology at Georgetown University. Her research focuses on musical form, music notation, and film music, and she plays piano. She is primary editor of *The Psychology of Music in Multimedia*, and appears in *Score: A Film Music Documentary*.

Peter Pfordresher is Professor of Psychology at SUNY Buffalo in New York State, USA. He completed his PhD in Psychology at the Ohio State University. His research on the relationship between perception and action in music has been published in psychology, music cognition, and music education journals, and has received support from the National Science Foundation. He sings and plays piano, guitar, and the trumpet.

Rom Harré is Emeritus Fellow of Linacre College, Oxford University, UK. He was Distinguished Research Professor of Psychology at Georgetown University, and Director of the Centre for the Philosophy of the Natural and Social Sciences at the London School of Economics.

Praise for the first edition

‘As it stands, Tan, Pfordresher, and Harré’s volume is an engaging exposition of the current state of our knowledge of the psychology of music …. To paraphrase Nietzsche, experiencing music without knowledge may not entirely be a mistake, but experiencing it with the kind of up-to-date knowledge that may be gleaned from *Psychology of Music: From Sound to Significance* is even more marvelous.’

– Aaron Kozbelt, *PsycCRITIQUES*

‘Tan et al.’s volume is an impressive achievement and merits serious consideration by anyone teaching a survey course in music cognition or seeking to recommend to a friend or colleague …. Its 300-odd pages are packed with the most detailed overview of our field we are likely to see in any text in the near term and it covers the major aspects of the field quite comprehensively.’

– Richard Ashley, *Music Perception*

‘[A]ll topics are introduced with sophistication and an interesting balance is provided between references to classical work and more recent work. Similarly, quantitative and qualitative investigations are both included …. The broadness of the book is splendid and allows for a complete introduction to the field. It will indeed be invaluable to undergraduate and postgraduate studies in the fields of psychology of music ….’

– Renee Timmers, *British Journal of Music Education*

‘The field of Music Psychology has received formative influences from many domains, thus it is no mean feat to create a representative survey of the literature. The authors have met the challenge, achieving a detailed and useful introduction to the field with this text.’

– Jessica Grahn, *Empirical Musicology Review*

‘We expect that this book will play an influential role in establishing the canonical organization for music psychology textbooks and hope to see it go through many editions in the years to come …. Meanwhile, we heartily welcome this ambitious book as a valuable new resource for teaching the psychology of music.’

– Roger Chaffin and Alexander Demos, *Psychomusicology*

Psychology of Music

From Sound to Significance

Second Edition

Siu-Lan Tan, Peter Pfordresher and Rom Harré

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

Second edition published 2018
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

and by Routledge
711 Third Avenue, New York, NY 10017

Routledge is an imprint of the Taylor & Francis Group, an informa business

First edition published by Psychology Press 2010

Cover design by Andrew Ward

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
A catalog record for this book has been requested

ISBN: 978-1-138-12466-0
ISBN: 978-1-138-12468-4
ISBN: 978-1-315-64802-6

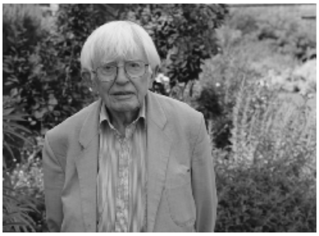
Typeset in Times New Roman
by Keystroke, Neville Lodge, Tettenhall, Wolverhampton

Contents

About the authors

Siu-Lan Tan is Professor of Psychology at Kalamazoo College in Michigan. She completed degrees in Music at Pacific Union College, graduate studies at Oxford University, and an MA and PhD in Psychology at Georgetown University, Washington, DC. She plays the piano and taught private lessons and music classes for many years in Hong Kong and California before shifting to Psychology. Her research focuses on listeners' perceptions of musical unity, graphic representations of music, and the role of music in film and other multimedia. Her work has been published in *Music Perception*, *Psychology of Music*, *Psychomusicology: Music, Mind, and Brain*, *Empirical Musicology Review*, and *Journal of Applied Developmental Psychology*, among other journals. She often speaks on the topic of the psychology of film music, and was the Keynote Speaker for the 10th *Music and Moving Image* conference at NYU. She is also primary editor of the book *The Psychology of Music in Multimedia*, and appears in *Score: A Film Music Documentary*, which was selected as a *New York Times* Critics' Pick. Currently, she serves on the Board of Directors for the Society of Music Perception and Cognition.

Peter Pfordresher's primary training is in experimental psychology, though he is also a classically trained pianist and vocalist, and was for years a semi-professional rock musician. He holds degrees in Psychology from Georgetown University (BA), University College London (MSc), and the Ohio State University (PhD). He is currently a Professor at the University at SUNY Buffalo, New York State. His research (supported by the National Science Foundation and the National Association for Music Education) focuses on the interplay between perception and action in music, individual differences in singing accuracy, retrieval errors in performance, and how the brain processes melodic and rhythmic information from melodies. This work has appeared in prominent journals such as *Journal of Experimental Psychology: Human Perception and Performance* and *Cognition*, music cognition outlets like *Music Perception* and *Psychomusicology: Music, Mind & Brain*, and music education outlets like *Journal of Research in Music Education*. He is an associate editor for the journals *Music Perception* and *Psychological Research*, and has served on the executive board of the Society for Music Perception and Cognition, as Secretary.



Rom Harré first graduated in mathematics and held a lectureship in the University of the Punjab at Lahore in Pakistan. After graduate work in philosophy at Oxford University, he turned to philosophy of science as University Lecturer at Oxford. This work led to studies in the philosophy of psychology and ultimately to theoretical and empirical work in the field. Harré is author or editor of over 65 scholarly books in numerous fields, and has held appointments as professor in many institutions around the world. He was Distinguished Research Professor of Psychology at Georgetown University in Washington DC, and Director of the Centre for the Philosophy of the Natural and Social Sciences at the London School of Economics. Rom Harré's books include *Pavlov's Dogs and Schrodinger's Cat* (Oxford, 2009), *Key Thinkers in 20th Century Psychology* (Sage, 2006), *Wittgenstein and Psychology* (Ashgate, 2005, with M. Tissaw), *Cognitive Science: A Philosophical Introduction* (Sage, 2002), and *One Thousand Years of Philosophy* (Blackwell, 2000).

Preface to the second edition

When instructors first prepare for a new course, they often focus on rolling out as much information as possible to provide the full topography of a field of study. After a few cycles of teaching the course, they begin to sense where explanations need more clarity, and can anticipate frequently asked questions, and provide more intriguing extensions to topics to which students are naturally drawn.

And so it is with our first and second editions of *Psychology of Music: From Sound to Significance*. The first edition was inspired by the need the authors saw for a text for a course which two of us had been teaching since the late 1990s, and one had begun to teach a decade later. It took a team effort of over four years to produce the first edition. Updates and expansions to the second edition have been largely shaped through dialogues with users of the first book, both instructors and students alike (including some of our former students who are now instructors for this course), along with our own experience teaching with the text in our courses. In addition, we actively sought feedback from our classes and student focus groups involved during the preparation of this edition (please see Acknowledgments).

While continuing to aim for the breadth of scope and level of detail of the first book, our second edition updates the content with developments in the field reflecting the seven years that have passed, and also warming the content and tone to connect the subject matter to our readers' natural interests and points of curiosity. As we continue to tackle questions that lie at the heart of our field, we also consider questions we have heard our students and other readers ask:

Why do I like listening to sad music? (ch. 14)
What does my friend mean when he says he's 'tone deaf'? (ch. 11)
Can you also be 'deaf' to the beat of music? (ch. 6)
How do you get rid of a tune that's stuck in your head? (ch. 5)
Can you train yourself to have 'perfect pitch'? (ch. 5)
Are some people just born to be more musical than others? (ch. 10)
What do babies understand about music? (chs. 8 and 9)
How do I deal with my anxiety about playing in front of people? (ch. 12)
How does my neighbor, who wears a cochlear implant, experience music? (ch. 3)

We also address many other meaningful questions.

Organization of the book

The organization of the book remains the same as the first edition, while chapters have been updated and often expanded with new topics. 'Part I: Foundations' explores the acoustics of sound, the auditory system, and responses to music in the brain. 'Part II: The Perception and Cognition of Music' focuses on how we process pitch, melody, meter, rhythm, and musical structure. 'Part III: Development, Learning, and Performance' describes how musical capacities and skills unfold, beginning before birth and extending to the advanced and expert musician. And finally, 'Part IV: The Meaning and Significance of Music' explores social, emotional, philosophical and cultural dimensions of music and meaning.

In keeping with the spirit of the first edition, we intend our book to engage a wide range of readers, and have kept technical terminology for music and psychology to a minimum to ensure the readability of the text for a broad audience. Only a general familiarity with music and psychology is assumed, and key words are defined and often exemplified. (The only exception is chapter 7, which discusses prominent music theories and inevitably requires more technical knowledge than any of the other chapters.) Our illustrative figures also do not rely heavily on musical notation, and when employed, notation is usually limited to a single-line melody. Keeping to the practice of the first edition, neuroscientific studies are mainly contained within our chapter on neuroscience and music, and in separate sections toward the end of our chapters on melody, rhythm, practice, performance, and emotion in music. Depending on the emphasis of the course or particular interests of the reader, this allows the option of reading chapters while omitting these sections, without disturbing the continuity of the narrative. Where neuroscientific findings are interspersed elsewhere in the text, they are described at a more general level suited to a broad readership.

What's new about the second edition?

Our second edition updates the content with developments in the seven years that have passed, integrating over 500 new citations and references to reflect advancements in the field. Many compelling new topics have been introduced as expansions to chapters. A summary of the significant changes to chapters is provided in Table 0.1 in the 'Notes to Instructors' following this preface.

While continuing to present material with the depth and detail that we hope will be a hallmark of this book in all its eventual editions, the technical content has been clarified with more vivid and meaningful descriptions, and more illustrative examples, especially in chapters 2, 3, 4, 5, and 6. Greater emphasis has been placed on exploring topics that apply the concepts of the book to real-world contexts, and to personally meaningful applications that will engage readers. For instance, the technical content on the topic of sound and hearing (chapter 3) is balanced with new sections focusing on the practical issue of music-induced hearing loss in musicians and in listeners using personal listening devices, and recent research on music perception among individuals who use cochlear implants. As noted earlier, many of these new sections and extensions were inspired by intriguing questions from our readers.

Further, the second edition incorporates additional resources for learners. First, in response to the suggestions of students and instructors, we have employed ***bold italics*** to draw attention to key terms, accompanied by clear definitions. In general, we reserve *regular italics* for conveying emphasis and for non-English terminology (e.g., *accelerando*). In addition, in chapters containing a lot of technical terminology (such as the sections on the ear and brain), more fine-grained terms appear in *regular italics*. Second, in addition to our 'Notes to Instructors,' we have included 19 exercises and practical application projects designed to promote active learning, and to cement the knowledge gleaned from each chapter. These are included in a new appendix, 'The Chapters in Action.' Some are individual or group assignments, and others are conceptual replications of key studies that can be carried out with minimal resources.

Finally, we have also devoted more attention to social facets of music, interspersing more links to the social realm in the two developmental chapters, the chapters on practice and performance, and a significantly expanded chapter on the social psychology of music. Although the book continues to focus on classical music as a model, reflecting the focus in the existing body of research in the field, we have included more jazz and pop music and other genres in studies and examples.

Personal note

The publication of this second edition will mark the 90th birthday of our co-author, Rom Harré. Thus he has passed the torch to the two of us. Revisions for this second edition were therefore carried out by Siu-Lan Tan and Peter Pfordresher. However, it is important to note that Rom's expansive knowledge as an eminent scholar in many fields, providing greater breadth of topics than the book would have otherwise encompassed, continues to be felt in the current volume and we therefore continue to include his name as an author.

We have expressed our gratitude to our supportive colleagues, students, and publisher in our Acknowledgments section. And finally – but of *foremost significance* in our lives – we thank those who were our brightest lights throughout the journey of both editions of this book: our families. Siu-Lan gives her heartfelt thanks to Danny Kim, Khoen, Nanny, and May-Lan Tan, and Brenda Kim. Peter extends special thanks to Lyn, Emma, and Paul Pfordresher for all of our musical times together.

Siu-Lan Tan, Michigan
Peter Pfordresher, New York
Rom Harré, Oxford

Notes to instructors

The second edition of this book reflects the growing number of courses being taught in psychology of music around the world, and the increasing interest in this subject reaching a wide readership. The authors are delighted to learn that this book is currently being used in a wide range of programs, including music cognition courses in psychology and music departments, required courses in music therapy programs, and special topic courses offered as electives for a variety of majors. The book also appears on required reading lists as a qualification for entry into some music psychology graduate programs.

While the growth and vibrancy of this field is certainly exciting, there are also challenges involved in producing a book intended to give a 'snapshot' of a vital and ever-changing field of research and to serve a diverse audience. Thus, we have done our best to design a book to accommodate courses with different emphases, and readers with varied backgrounds in psychology and music. Early chapters provide general explanations of basic musical terms and concepts such as 'interval' and 'key' (chapter 5) and 'beat' and 'meter' (chapter 6), and we have expanded this basic coverage in the second edition. Throughout, key terms relating to psychology or the specific topics at hand are defined or explained, and visually accented by ***bold italics***. As mentioned in the preface, the structure of the book offers the option of reading chapters without the neuroscientific sections toward the end of some of the chapters. Readers will still gain some familiarity with findings of studies focusing on the brain, as they are also woven into other parts of the text at a level that is suitable for a broad readership.

Also new to this edition is our appendix, entitled 'The Chapters in Action.' In response to requests we have received from fellow instructors since the first edition, we offer 19 brief exercises and more extended application assignments or projects that are presented in a flexible 'accordion-style' manner – adaptable to different class sizes and academic levels, individual or group projects, and courses with different emphases.

No single psychology of music book can cover all territories any longer, thanks to the vibrant growth of this field of research since the days when the focus was almost solely on perception and cognition of music. Today, the psychological study of music has expanded in developmental, social, neuroscientific, and applied directions, and numerous subspecialties. Our book is designed to give readers a firm grounding in some of the main domains of the rich and interdisciplinary field of psychology of music, to be supplemented by primary readings selected by instructors to extend to their own research topics and areas of interest, and to motivate readers to continue exploring topics that pique their curiosity.

Changes since the first edition

We recommend that instructors see the Preface for a broad overview of the book and some new features of the second edition. The organization of the book and internal structure of most chapters remain the same, although with many new topics and expansions. Chapters 5 and 14 have undergone the greatest changes due to the growth in these areas, but the chapters maintain the same general structure. The main changes and additions since the first edition are summarized in Table 0.1.

Table 0.1 Overview of changes since the first edition.

1	Introduction	Updated with new examples. Otherwise, this brief opening chapter remains much the same.
2	Acoustics	Improved clarity of discussion on acoustical properties of sound through examples relating these properties to music. Expanded discussion of Fourier analyses to emphasize relevance to music. Updated studies on acoustics of concert halls, opera houses, and other musical venues.
3	Hearing	Improved clarity of technical content on perceptual qualities of tone and physiology of the ear. New section added on music-induced hearing loss in both musicians and listeners, including risks involved in the use of personal listening devices. New section added on music perception in individuals with cochlear implants.
4	Neuroscience	Reorganized to integrate discussion of research on music and language in the brain, including the question of whether music and language constitute separate neural 'modules.' New discussions of diffusion tensor imaging and neural plasticity as related to the effects of musical training.
5	Melody	This chapter has been overhauled. Expanded discussion of basic musical concepts with application to musical examples, including the relationship between pitch height and chroma, and the distinction between absolute and relative pitch. Three new figures help illustrate these concepts. Expanded coverage of absolute pitch perception. New coverage of 'earworms.'
6	Rhythm	Revised to describe basic concepts more clearly. Updated survey of research includes beat deafness and the perception of 'groove.'
7	Structure	Clearer description of how music may be said to have a 'grammar.' Otherwise, the chapter remains much the same.
8	Development	Updated and expanded sections, especially on infant perception of pitch, meter, and memory. Inclusion of infant music cognition studies with social implications. Expanded section on infant-directed singing coordinating musical and social synchrony between caregivers and infants, including brief examples of therapeutic use of maternal voice and music with infants.
9	Learning	Expanded sections on development of singing and movement to music. Musical play topic extended to older children, including musical games in playgrounds. New section on children's composition, both solo and collaborative. More attention to the role of technology, in musical play and computer-assisted composition.
10	Practice	Expanded discussion of the role of 'talent,' genetics, and deliberate practice. Expanded section on informal practice. New section added on injuries and other conditions prevalent in musicians, particularly playing-related musculoskeletal disorders, and passage on focal dystonia added to section on music practice and the brain. Updated studies on practice strategies.
11	Performance	Increased use of examples to promote reader engagement. Discussion of new studies concerning the role of memory in improvisation, motion capture analyses of performance, and poor-pitch singing. Includes discussion of emotional communication in performance, previously in chapter 14.
12	Social psychology	New section on music performance anxiety and complementary topic of performance boost. Expanded section on conductors, including effects on performers and audience. Expanded section on gender and music, including gender stereotyping of musical instruments and genres. Greater breadth on music and consumer behavior, including discussion of possible mechanisms.

13 Meaning	Added new citations of research relevant to the conveyance of meaning in music. Otherwise, the chapter remains much the same.
14 Emotion	This chapter has been overhauled. Expanded discussion of research on both perceived and felt emotions, and broadened scope of emotions being investigated. New section added on the use of music for emotion regulation, including mood maintenance, enhancement, and rumination. Brief section added on listener engagement with music evoking sadness. Updated section on film music. (The passage on how performers express emotion was moved to Chapter 11. The passage on emotion and culture was moved to Chapter 15, where it has been expanded).
15 Culture	Updated discussion of musical abilities in nonhuman animals and emotional communication in music. New sections on how non-Westerners perceive Western music, and memory for music of other cultures.

Exercises and application assignments: ‘The Chapters in Action’

For exercises and application assignments designed to accompany the chapters in this book, please see our new appendix, ‘The Chapters in Action,’ which can be found toward the back of this book. It includes 19 brief exercises and application assignments designed to involve students in active learning, in tandem with chapter reading. We hope that these ideas will serve as a starting base that instructors can tailor to fit their classes, and will inspire further creative ideas to make the chapters come to life for students.

Our goal was to offer a book of broad scope, to provide instructors with a range of topics to choose from, to tailor to the level and composition of the class, and to meet the goals of different courses. It is our hope that this book will continue to serve a wide array of courses and draw many more inquisitive, bright minds to this fascinating field of study, as there is still much to explore and discover.

Acknowledgments

It takes a village to write a book, and no less of a community to create the second edition. We are grateful to many colleagues and students who contributed so generously of their time and talent to this book.

The authors would like to thank Kate Gee, Timothy Justus, James Mantell, and Douglas Vipond for offering particularly instructive suggestions for our second edition. We also drew inspiration from insightful published commentaries by Richard (Ric) Ashley, Roger Chaffin, Alexander Demos, Aaron Kozbelt, Jessica Grahn, and Renee Timmers. The authors are also grateful to our colleagues Arianne Abela, Gillian Anderson, Jonathan Ashmore, Leo Beranek, Meagan Curtis, Autumn Hostetter, Timothy Justus, Andrew King, Josh McDermott, Israel Nelken, Caroline Palmer, Jessica Phillips-Silver, Christopher Plack, Jan Schnupp, Michael Schutz, Mark and Cheri Shevy, and Aaron Williamon for their generous contributions to the second edition. We also thank Kathryn Lightcap, Graphic and Multimedia Designer, for expert assistance with figures.

As with the first edition, valuable student input shaped this book from its earliest drafts to the final manuscript. We express special appreciation to Elizabeth (Elie) Penix, Mackenzie Norman, Patrick McGuire, and Karen Chow for particularly detailed suggestions and meticulous editing of selected chapters. We thank our many dedicated student assistants for valuable feedback and assistance: Christina Dandar, Ashley Schmidt, Camila Trefftz, Emma Greenspon, Jacylyn Stoebe, Tim Pruitt, Sarah Teh, Amanda Knose, Jessica Troung, Nicholas Horwood, Haley Haner, Paul Kovacs, Anthony Khoury, Ryan Rondinaro, Anthony Nagib, Tiffany Thavisin, Thomas Gadelrab, and students in our psychology of music classes.

We are grateful to the wonderful staff at Routledge, Eleanor Reedy, Michael Strang, Elizabeth Rankin, Alex Howard, and Libby Volke, for all their guidance and support and for being genuinely caring advisors to us. We would also like to thank our dedicated copy-editor Huw Jones, Kelly Winter and the Keystroke staff, and indexer Michael Hamilton, for all the care and attention they put into the preparation of this book. And finally, our special gratitude goes to publisher Lucy Kennedy at Routledge, who first came across our original manuscript and advocated for the publication of our book.

From the first edition, we continue to extend appreciation to our colleagues Leo Beranek, Steven Brown, Ginevra Castellano, Roger Chaffin, Jennifer Cox, Simone Dalla Bella, David Hargreaves, Paul Jeffries, Patrik Juslin, Timothy Justus, Rohan Krishnamurthy, Scott Lipscomb, Martin Lotze, Josh McDermott, Rod Nave, Jessica Phillips-Silver, Christopher Plack, Dirk-Jan Povel, Mari Riess Jones, Michael Schutz, Keith Swanwick, David Temperley, Michael Tenzer, Laurel Trainor, and Robert Zivadinov; and to our former students (many of whom are now professors and researchers, and some who have taught courses with this book), Matthew Bezdek, Ryan Coppolo, Emily Dayton, Timothy Griffiths, Jackie Howard, John Kulpa, James Mantell, Amanda McFarland, Lauren Mitchell, Christy Peaslee, Christina Violante, Shanti Virupannavar, Elizabeth Wakefield, Sally Warner Read, and to all our psychology of music classes; and to the dedicated staff at Taylor & Francis who guided the first edition: Lucy Kennedy, Tara Stebnicky, Sharla Plant, Dawn Harris, Nicola Ravenscroft, and our copy-editor Paula Clarke. Our gratitude to all.

1 The scope of psychology of music

People everywhere and at all times for which we have records have picked out certain patterns of sound for particular attention. Some of these patterns are the stuff of what we call 'music.' What are the characteristics of the sound patterns we recognize as music? Why is it that these sound patterns have had a special significance for human beings?

All perceptible sounds begin with a propagation of energy into the environment. It may be a light breeze setting into motion a thousand fluttering leaves, the plucking of the strings of a harp, or the striking of a bass drum. What makes the particular dance of air molecules 'musical' in some instances, while other disturbances of air molecules seem to give rise to mere sounds? Or noise?

It is not always possible or desirable to give a formal definition of the topic of a program of study. We will not try to answer the question 'What is music?' in a neat, short formula. There are paradigm cases of music that most people in a particular culture can recognize. We can start with some exemplary cases from our own culture, such as a symphonic performance, a rock concert, or an advertising jingle. Many hear church bells, ringing out a simple melody, as music. However, not all sound is music. Could we draw up a similar catalogue of sounds that everyone would agree are not music? Perhaps the sounds of busy city traffic or the whine of a vacuum cleaner might strike us as obvious cases of nonmusic. But what about the sound of waves breaking on the beach? The howl of a wolf? Or the song of a bird?

While the extremes seem clearly distinguishable, there is no sharp line to be drawn between music and sounds that are not music. Whether sounds are taken as music or not depends in part on the context in which they occur. Debussy imitated the sounds of breaking waves in *La Mer*. Respighi's orchestral score for *The Pines of Rome* included a recording of a real nightingale's song. The blasts of real car horns punctuate Gershwin's energetic orchestration in *An American in Paris*, while Edgard Varèse – who questioned the distinction between 'music' and 'noise' – included two hand-cranked fire engine sirens in his influential percussion piece, *Ionisation*. And Malcolm Arnold's *Grand, Grand Overture* included three vacuum cleaners among the orchestral instruments!

In Hollywood films, the score often doubles or produces some of the sound effects accompanying a scene. In line with the idea that in space 'there is nothing to carry sound,' Steven Price's Oscar-winning score for the 2013 film *Gravity* straddled both music and sound design, providing a soundscape that conveyed the expansiveness of space, the tension and emotion of each scene – and afforded something akin to sound effects, enabling the audience to feel the impact of the shuttle breaking apart and the massive explosions that sent debris flying through space. There is also John Cage's *4′ 33″*, a composition in which the performer does nothing for 4 minutes and 33 seconds to allow ambient and incidental sounds to define the composition. Here, the boundaries between sound and music, and the very definition of music in the absence of sound controlled by the composer or musician, are brought into question.

All-encompassing definitions of music may be elusive. Nevertheless, despite the fuzzy boundaries of the domain of music, it seems that there are auditory phenomena which we can generally agree are music, and on which there has been agreement in many cultures and historical epochs. Music is produced and perceived by human beings. Performers must learn the necessary skills to create ordered sound in meaningful patterns; through exposure or training, listeners must learn to perceive those features of ordered sound patterns as music. There is clearly room for a systematic study of all these skills, as diverse as they may be. The merging of psychology and music leads the way to such examination, and opens avenues to numerous and diverse topics for study.

The scope of the field

The psychology of music is motivated by a great many questions. Among other things, the field of ***psychology of music*** is concerned with *the processes by which people perceive, respond to, and create music, and how they integrate it into their lives*. These topics range from the way in which the ear extracts the pitch of a tone, to the way in which music is used to express certain emotions or transform moods. Though this field of study makes important use of *cognitive psychology*, it also draws on many other branches of psychology such as *sensation and perception*, *cognitive neuroscience*, *developmental psychology*, *social psychology*, and applied fields such as *educational psychology*.

The perspectives of each of these domains within psychology and more have shaped this book to some degree, as evident in the overall plan: Our exploration begins with a consideration of the physical properties of a sound wave, the transmission of sounds to the ear, and the neural bases of the perception and cognition of music. We then examine more closely the perception and cognition of melody, rhythm, and musical structure. Next, we trace the emergence and development of auditory capacities and musical abilities, and the acquisition of musical expertise, culminating in musical performance. Finally, we consider the question of meaning in music, and the social, emotional, and universal significance of music.

Interdisciplinary connections

The psychology of music attracts not only psychologists and musicians, but scholars and researchers from a wide range of other disciplines. The present volume is also informed by perspectives from fields such as *acoustics*, *neuroscience*, *musicology*, *education*, *philosophy*, and *ethnomusicology*, among other disciplines. We will provide a brief overview of these connections here.

More than 2000 years ago it was realized that the musical possibilities of sound as heard were shaped and constrained by the physical properties of sound waves as they interacted with the amazing powers of the ear. Pythagoras linked the weight of a vibrating object to pitch, while his followers extended his intuitions to include the vibrations of strings, linking pitch to the length of the string, and harmonics to the simple numerical ratios of those lengths. *Acoustics* is the science of the production, propagation, and reception of those vibrations in the air that are relevant to hearing in general and music in particular. But it is also more – as discussed in chapter 2, the way that musical instruments and the human voice shape the physical processes that reach the listener, as well as the properties of the performance venues where music is produced and enjoyed, are also parts of the science of acoustics.

With respect to *neuroscience*, recently there has been a real surge of interest in the neural underpinnings of human musicality, and the current volume considers the way that neural activity may constrain or enhance our experience of music and music-making. In order for us to experience music, the brain must pick up and import physical patterns from the auditory signal. Our studies must include an introduction to auditory neuroscience, beginning with the anatomy and physiology of the ear as presented in chapter 3, and moving on to a comparison between the neural bases of music and language processing in chapter 4. Recent discoveries from the field of neuroscience of music are also presented in sections within subsequent chapters on perception of melody and rhythm, music practice and performance, and emotion in music in chapters 5, 6, 10, 11, and 14 respectively, and sprinkled elsewhere in the book.

Insights from the field of *musicology* (the study of the structure and history of music) also continue to be essential to psychology of music, as evident in chapters 5, 6, and 7 on perception of melody, rhythm, and musical structure, and in chapter 14 on the emotional power of music. For example, in studying the power of music to 'express' emotions and its capacity to 'induce' emotional states in the listener, musicologists have addressed intriguing questions such as: How do musical structures give rise to emotions in listeners? Further, how does music come to have 'meaning' for the listener? The latter question is explored in chapter 13. A study of the basic theory of a practice involves bringing at least some of the underlying presuppositions to light as explicit principles, and subjecting them to critical examination. This is one of the ways that *philosophical* analysis is also an indispensable part of the psychology of music, bringing out certain presuppositions in the practices by which psychologists try to reach an understanding of music as a human phenomenon.

Figure 1.1 Infant music cognition is just one of many growing areas of research that reflects the expansive scope of study in this field. Investigations in this area touch on various domains of study, such as sensation and perception, cognition, neuroscience, developmental psychology, social psychology, and applied areas such as music education.

Source: Photograph used with permission from the parent. Copyright © Jessica Phillips-Silver.

Within psychology of music, developmental psychologists are concerned with the emergence and maturation of musical behaviors, some of which are described in chapter 8 on music perception and cognition in infancy. How these emerging abilities may best be supported and refined is a question for *music education*, the focus of chapter 9. Musical performance requires the development of a set of highly elaborated skills, and a growing base of knowledge that allows for the sensitive interpretation of music. Innovative music education methods assume that all children are musical, and intensively immerse young children in creative and sensitive engagement with music with the aim of laying the foundations for lifelong musicality.

Finally, music is differentiated into a wide variety of musical cultures across the world, and as such, an anthropological perspective highlighting the study of distinct human cultures and their 'musics' is also relevant to an understanding of the psychology of music. These studies help us to differentiate between cultural sources of the features of a particular musical repertoire and those that may derive from the biological bases of musical perception. The application of anthropology to music is *ethnomusicology*, a major component of chapter 15. In this final chapter, we consider music as it is represented in a variety of societies around the world. But we shall not pretend to have answered entirely satisfactorily the enduring questions of what music *is*, how exactly it is created and received, and how music brings about the powerful effects that it does.

Range of research methods

Although the present discussion by no means exhausts all the questions and topics subsumed by the study of psychology of music, the richness and expansiveness of this field should already be apparent. Psychology of music is also distinguished by the scientific research methods commonly employed to study musical phenomena of interest, which leads one to a certain kind of explanation for the phenomena under investigation. There are many different kinds of explanations developed in the natural sciences, accompanied by various methods, and psychologists of music make use of many of them.

Throughout this book we will encounter a great variety of empirical studies, as various methods are needed to explore the whole gamut of musical experience. For some purposes, experiments manipulating specific variables in carefully controlled conditions are useful. A researcher may manipulate the tempo (or pace) of a song to determine if it alters listeners' interpretations of the emotion that is being expressed. For other purposes, the recording and analyzing of real-time phenomena of musical production and experience are appropriate, such as when identifying the steps a concert pianist takes to learn and memorize a complex piece of music. In some instances, physiological methods or brain-imaging techniques such as functional magnetic resonance imaging (described in chapter 4) may be employed, for example to examine changes in the body or brain while listening to or performing music. Some investigators have even monitored the heart rate, breathing rate, and brain activity of musicians and audiences during live or virtual performances. Methods such as motion-capture may be used to track the fine motor movements of a musicians' actions in real time, or to see how a musical quartet may coordinate their collaborative performance with communicative gestures. Electromyography may be used to measure electrical activity of muscles, in order to monitor subtle facial muscle movements as a singer is watching another vocalist perform. These are all examples of research topics discussed in the present book.

Qualitative studies and naturalistic observation also play an important part in studying the psychology of music. The nature of some aspects of musicality is best captured in rich observational studies of spontaneous activities in natural settings, such as observations of crowd behavior at live concerts, or of children playing musical games in playgrounds, or of parents singing to infants in their homes. These topics of study are also described in this book. Sometimes we can usefully sum up the musical experiences of a great many people in a sweeping generalization. In other cases, we need to pay close attention to music as it is perceived, appreciated, performed, or composed by an individual, if we are to reach a sufficiently detailed understanding of what is happening.

Since the days of the pioneering volumes in this field, including Carl Seashore's *The Psychology of Musical Talent* (1919) and *The Psychology of Music* (1938), Paul Farnsworth's *The Social Psychology of Music* (1958), Geza Révész's *Introduction to the Psychology of Music* (1954), and Robert Francès' *The Perception of Music* (1958/1988), a wide variety of methods have been applied to the study of psychology of music. It is a great mistake in building up a science to insist on the primacy of any one of the many methods of research that can be found across the universe of the sciences. Indeed, the psychology of music is fascinating in part because it draws on so many diverse bodies of knowledge and multiple modes of investigation! Only by adopting such breadth can we hope to understand how this remarkable human practice is possible.

Coda

Our introductory chapter has mapped the rich interdisciplinary terrain of the present volume. This book was undertaken by authors who first met some 25 years ago, when their paths first crossed. Brought together many years later by a shared passion for this field, we collaborated on this volume in response to our many enthusiastic students who have influenced us during almost two decades of teaching psychology of music classes. Clearly, the scope of inquiry of this field is vast, and no book can cover all that psychology of music encompasses. Our aim is simply to introduce readers to a broad range of classic and current research, and to ignite curiosity in the many intriguing questions and topics in this exciting and expanding field.

Considering what music meant to him after 50 years of studying the science of music, Carl Seashore – one of the important pioneers in this field – wrote in the preface to his 1938 volume of *Psychology of Music*:

> Then I was a stargazer; now I am an astronomer. Then the youth felt the power of music and gave expression to this feeling in the way he loved and wondered at the stars before he had studied astronomy. Now the old man feels the same 'power of music,' but thinks of it in the manner that the astronomer thinks of the starry heavens.
>
> (p. xi)

Like Seashore, we hope this volume will serve to bridge the gap between the 'mere love and practice of music to an intelligent conception of it' (p. xi). At the same time, while learning the various principles and practices of astronomy, we hope our readers continue to be stargazers – moved by music and constantly fascinated by its many enduring mysteries.

Recommended resources: Psychology of music

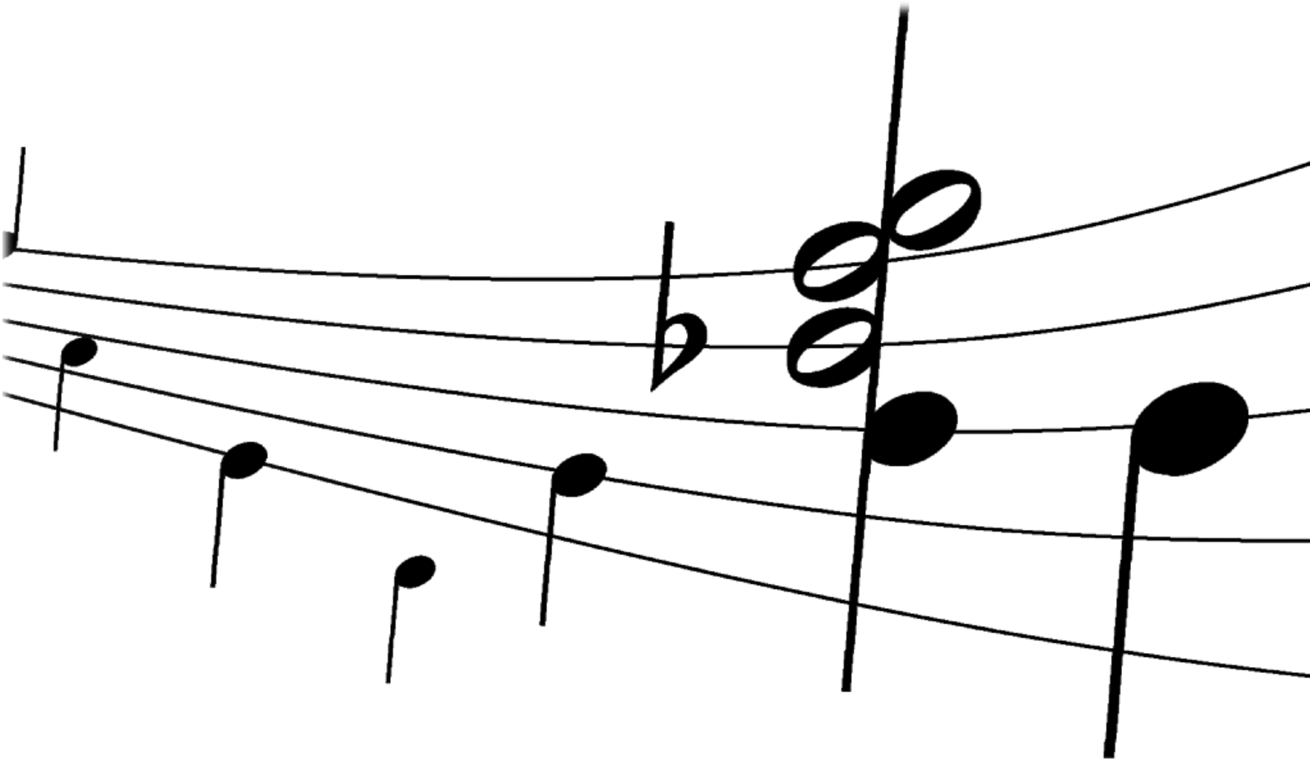
Ashley, R., & Timmers, R. (Eds.). (2017). *The Routledge companion to music cognition.* New York: Routledge.

Thompson, W. F. (2015). *Music, thought, and feeling: Understanding the psychology of music* (2nd ed.). New York: Oxford University Press.

Patel, A. D. (2015). *Music and the brain*. Chantilly, VA: The Great Courses [eighteen 30-minute lectures on topics complementing subjects in this book, available in video or audio format].

Part I

Foundations



2 The acoustics of music

Brittle twigs and dry leaves crunching underfoot, rain pellets hammering down on a tin roof, the roar of a crowd at a football game. Each of these sounds has its characteristic acoustic features and patterns that make it distinct from others. A distant flute on a hillside, the shimmering tones of a *gamelan* ensemble in a garden courtyard, the thunderous *fortissimo* chord of a full symphonic orchestra in a grand concert hall. Each of these musical sounds, too, has a distinct acoustic signature. All these auditory events – musical and nonmusical – are dispersed into various sound environments: open air, outdoor arenas, and crowded concert halls. The sound waves often travel a great distance to our ears, and the physical patterns must be transformed into neural signals in the ear and brain. It sounds like a complex process. Yet most of us quite effortlessly recognize the sounds and their sources, defining some as 'noise' and some as 'music,' and describing them as 'sweet' or 'rough,' 'mellow' or 'shrill.'

Most of this book focuses on the subjective experience of music as a result of neural, cognitive, developmental, and social variables. However, the raw material that makes up music is the physical stimulus we call sound. Note that this stimulus is something distinct from either the object making the sound (e.g., a guitar), or our ear's response to the sound. ***Acoustics*** is the science of sound, referring specifically to its production, transmission, and reception (*American Heritage Dictionary of the English Language*, 2016). The particular domain of *musical acoustics* focuses on the 'mechanisms of sound production by musical instruments, effects of reproduction processes or room design on musical sounds, [and] human perception of sound as music' (Hall, 2001, p. 2).

The present discussion begins with a brief description of the *physical* characteristics of sound waves (*frequency*, *amplitude*, *power spectrum*). We then illustrate a few principles of acoustics as they apply to a variety of musical instruments. Finally, we explore the acoustics of some grand concert halls and opera houses around the world. The perception of sound as music and the *perceptual* dimensions of sound (*pitch*, *loudness*, *timbre*) will be further discussed in chapter 3, accompanying a description of the auditory system.

Sound as physical stimulus

When a guitarist plucks a string, the resulting vibrations of the string lead to a corresponding disturbance in the air surrounding it. Air is important here; it functions as a medium through which sound propagates. As we will discuss later, the hollow body of an acoustic guitar is also important in amplifying the initial vibrations in the air that surround the guitar string. These physical disturbances propagate through the air and thus form a ***pressure wave***. Without some density in the air molecules, sound cannot travel and the environment is silent.

For instance, although sounds associated with the engines of giant spaceships as they streak through deep space makes for great dramatic effect in movies like *Star Wars*, in reality these battles in space would be entirely silent!

Figure 2.1 illustrates sound propagation. It shows a tuning fork, which creates sound similarly to plucking a guitar string. This figure allows us to explore the structure of a pressure wave in more detail. When the tuning fork is struck, its tines are set in motion. As a tine moves in an outward direction, the air molecules directly adjacent to it cluster together, leading to ***compression***. Then as the tine moves back in past its midpoint, the molecules spread apart, leading to ***expansion***. Of course, the tuning fork's tines, like a guitar's string, move back and forth (oscillate) rapidly over a period of time. As a comparison, think of what happens when you move your arm through the water in a swimming pool. As you move you arm to the right, your movement creates a bulge in the water that you push against. This is compression. Behind your arm, there is a kind of valley in the water – expansion.

This oscillating process leads to alternations of compression and expansion in the pressure wave surrounding the tine, and these disturbances ***propagate*** (i.e., spread out) across the surrounding air molecules. Thus, over time the pattern of compression and expansion leads to a ***sound wave*** (an acoustic pressure wave) that travels from the source of the disturbance (the tine or the guitar string) all the way to our ears, which respond to the sound wave in a manner described in chapter 3. In Figure 2.1, regions of compression are shown as dark bands to represent the clustering together of air molecules, and the spreading of molecules for regions of expansion is shown as light bands.

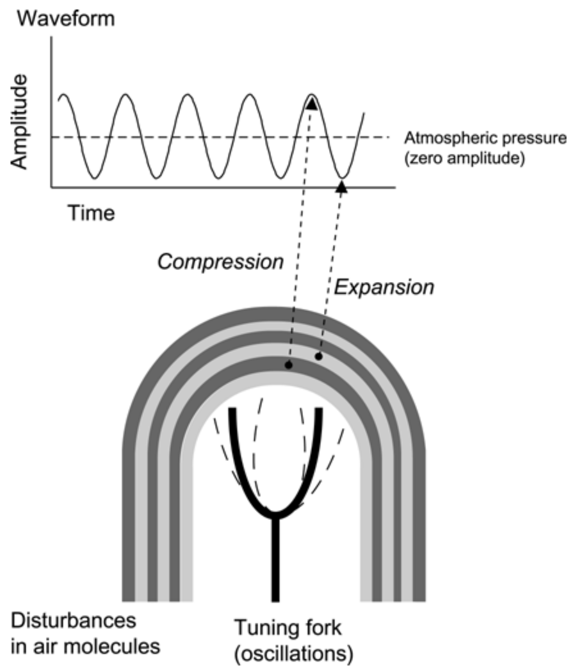


Figure 2.1 Propagation and representation of sound. Movements of the tines of a tuning fork (bottom) lead to alternating bands of compression and expansion among air molecules. For a pure tone (as in a tuning fork), the pattern of disturbances across time (for a given point in space) can be represented as a sine wave (shown at the top).

How does a sound wave propagate in order to reach our ears? A useful (though imperfect) analogy is to think again of water. When you throw a stone in a pond, the stone's impact creates ripples (compression and expansion) that radiate out from the source of its impact, though the extent of these waves diminishes with increasing distance from the source. Like a sound wave, the movement of these oscillations is longitudinal. In any ***longitudinal wave***, the movement of the oscillation is parallel to the direction of movement overall (as opposed to a 'transverse' wave, characteristic of light, in which the oscillation is perpendicular to the overall motion).

The movement of longitudinal waves can be demonstrated with a spring or Slinky toy. When a momentary force is applied to the first coil, it nudges the coil adjacent to it, which in turn nudges the next coil, and so on. The disturbance travels through to the other end like a wave, but the coils themselves do not travel from end to end. Each coil of a spring makes an oscillating motion that contributes to the wave and seems to pass it along, and then returns back to its place. In a sound wave, molecular movement affects adjacent molecules so that a periodic pressure wave is propagated through the medium (i.e., air) – but the constituents of the medium (i.e., the molecules) move only locally.

Although the essence of sound is a disturbance, a single discrete acoustic disturbance – such as a book falling on the floor – will not lead to the experience of music. Rather, the disturbances that lead to music follow regular patterns, resulting from vibrations in some sound-producing source. This is true of the plucked guitar string. Although the physical gesture is discrete (a single plucking motion), the string's vibration will continue for some time until the vibrations either diminish or the sound is stopped by another physical motion (e.g., damping the string with your hand). When the physical vibration is sustained, the surrounding air molecules begin to vibrate and the disturbance travels away from the source of sound, in an ever-widening wave of compressions and expansions, as shown in Figure 2.1. As it turns out, patterns of vibrations are integral to the link between sound and music. In the section that follows, we discuss different properties of these vibrations, and the sound waves (ultimately music) that result from them.

The properties of sine waves

One reason we used a tuning fork in Figure 2.1 rather than a guitar is that the tuning fork's vibration pattern is of the simplest sort: a sine wave (also called a 'pure tone'). A ***sine wave*** has a simple defining feature: Its pattern can be described fully using one frequency of vibration, which comes from the timing of alternation between compression and expansion phases. This vibration takes the smooth curved shape of a sinusoid because this is the natural pattern that a mass-spring system follows, and sound events fit this general model. Aside from the tuning fork, hardly any other sounds we hear in the real world can be fully described as a pure tone. Nevertheless, a sine wave is a clear way of representing important characteristics of sounds. Two parameters (i.e., measurable features) of sine waves that are critical for our discussion of musical sound are illustrated in Figure 2.2.

One parameter has to do with the rate at which a sine wave alternates between compression and expansion: ***frequency (f)***, the rate at which the crests or troughs of the wave pass a point in a given measure of time. Every time the sine wave completes a full cycle from compression to expansion and back again is considered a single period or 'cycle.' The number of cycles per second is referred to as ***hertz (Hz)***. Thus a sound source that is vibrating at 50 cycles per second is said to have a frequency of 50 Hz. In Figure 2.2, sine waves plotted in the top row have a lower frequency (1 cycle per second) than those in the bottom row (2 cycles per second; note that both signals would be below the frequency range of human hearing: see chapter 3). Many

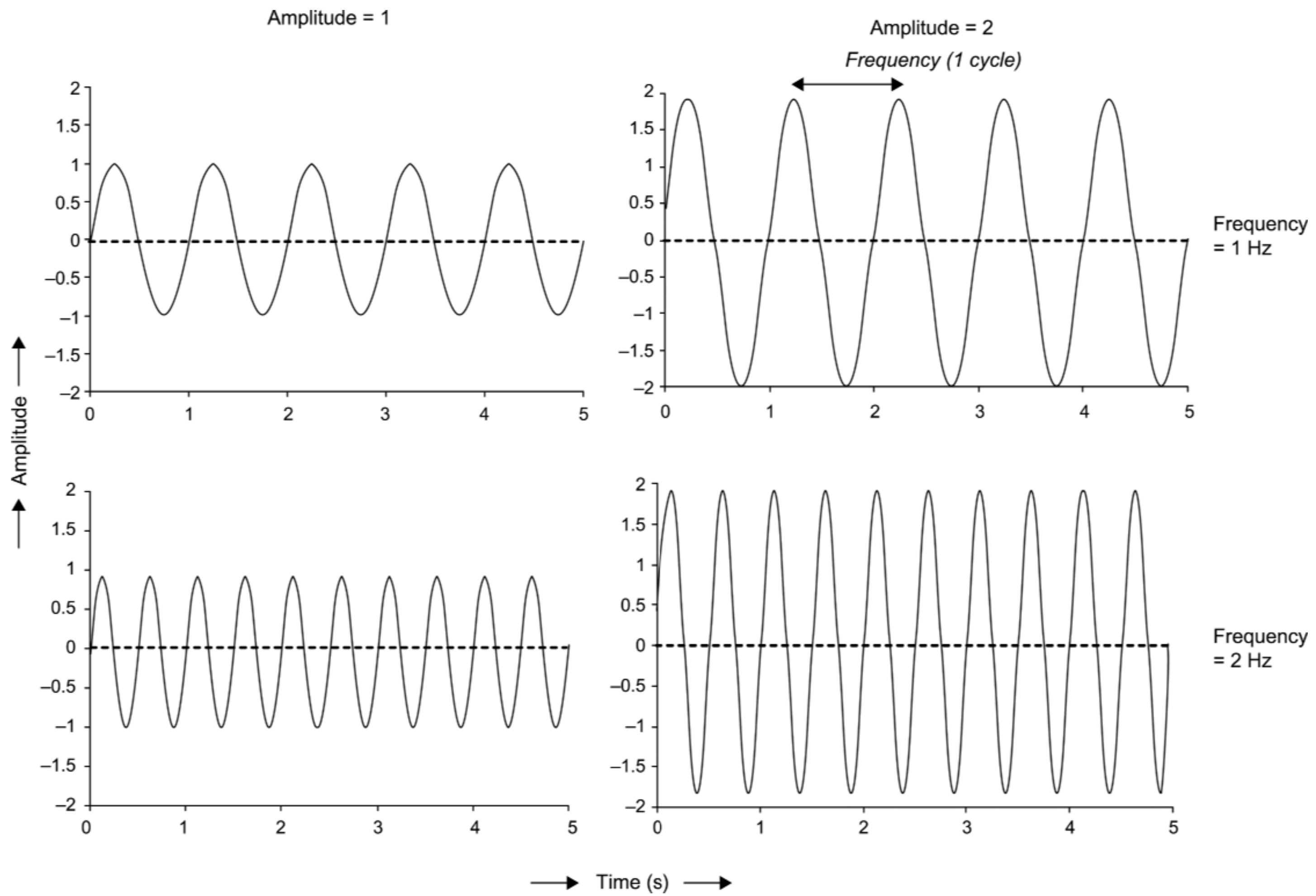


Figure 2.2 Four sine waves, varying in frequency and amplitude.

Source: Copyright © Peter Pfordresher.

physical variables can control frequency; a particularly important one is tension. When you pluck an open string on the guitar, then turn the tuning knob to tighten the string, and then play the same string again, you will find the pitch on the second pluck is higher than the first because the tuning adjustment has increased the tension in the string. A similar phenomenon happens in brass instruments when one tightens the lips (embouchure) while making the lips buzz, or when singers increase pitch by tightening the vocal folds (see chapter 11). Increasing tension causes sound frequency to increase. This in turn influences a perceptual variable: pitch.

In 1563 the mathematician-physicist Giovanni Battista Benedetti discovered that the frequency of vibration of the source of a sound gives rise to the aural sensation we call ***pitch***. We will discuss pitch at length in chapters 3 and 5; for now it is enough to define pitch as related to the experience of sounds as varying from high (a soprano singer) to low (a bass). When the frequency is high, there are many cycles in a given span of time, and the sound is perceived as high-pitched. A tone with a low frequency has few cycles in a given span of time, and the sound is perceived as low-pitched. Extending the physical to the musical, consider the fact that changes in this simple parameter, frequency, can enable us to experience the opening phrases of 'Mary Had a Little Lamb' or 'Twinkle, Twinkle, Little Star' as distinct melodies. It is important to note that pitch, though related to the physical frequencies of sounds, refers to subjective experience, and not to physics. Thus, whereas frequency refers to the physical signal, pitch relates to our experience of frequency information. We elaborate more on this important distinction in chapter 3.

The second parameter of a sine wave that is relevant to a discussion of music is ***amplitude***. The amplitude of a wave is the maximum displacement compared to the resting state, and corresponds to the perceptual dimension of sound that we experience as ***loudness***. Generally, the greater the amplitude of a wave, the more energy it transmits. Sound waves with more energy are generally perceived to be louder than sound waves with less energy, though the subjective effect depends to some extent on the frequency of the sound. In Figure 2.2, sine waves on the left side have lower amplitudes than those on the right side. In terms of physical gestures that lead to sound, the amount of force involved in a gesture roughly leads to changes in amplitude of the resulting vibration, which in turn affects the subjective experience of loudness. On the guitar string, a more forceful plucking action will lead to a greater initial displacement of the string, and thus a larger magnitude of alternating compression and expansion in the surrounding air.

Complex sound waves

Sine waves, as shown in Figures 2.1 and 2.2, have only a single frequency of vibration, which is why they are often called 'pure' tones (just as light with only one wavelength is called a pure color). But, as we will see, most sounds are best understood as a combination of many frequencies that occur simultaneously. The pure tone produced by tuning forks, beeps produced by automated teller machines, and old computers are well described by sine waves, but a friend's voice, the sultry saxophone, and other sounds of nature require a more complex description. Even a single plucked guitar string does not produce a sound wave that looks like a sinusoidal alternation between compression and expansion. Although the resulting sound wave may have a recognizable recurring pattern, it will not be sinusoidal and not be reducible to a single frequency of vibration. Not surprisingly, such sounds are referred to as ***complex tones***. An example is shown at the top of Figure 2.3.

If most natural sounds are complex, why discuss sine waves at all? The sine wave formula ends up being critical in order to make sense of the physical structure in a complex tone.

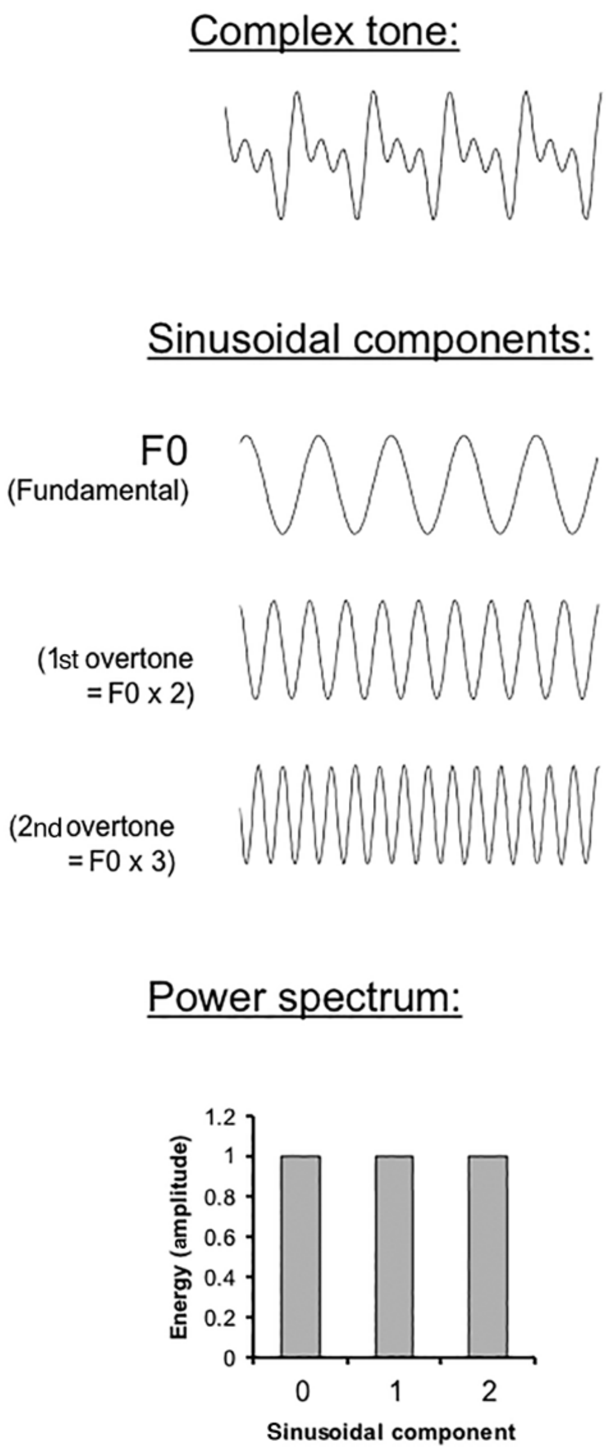


Figure 2.3 A complex wave, and below it a Fourier analysis.

Source: Copyright © Peter Pfordresher.

Specifically, any complex tone can be understood as a combination of sine waves. The process by which a complex wave is decomposed (broken up) into a set of component sinusoids is referred to as ***Fourier analysis***, named after French mathematician and physicist Jean-Baptiste-Joseph Fourier (1768–1830), who developed this technique.

Figure 2.3 illustrates how a complex tone can be decomposed into a series of sine waves. In this case, we need three sine waves (or 'components'), which are shown below the complex tone. The sum of these components yields the complex tone. Although the Fourier analysis was originally developed with engineering applications in mind, its application to sound does yield some important psychological implications because different components of a complex tone play different roles in perception. The component with the lowest frequency (longest wavelength) is commonly referred to as the ***fundamental frequency***. This frequency is often (though not always, as described in chapter 3) associated with the pitch of a complex tone. Higher-frequency components are often referred to as ***overtones***. The pattern formed by the overtones, called the overtone series, leads to a quality of sound referred to as ***timbre*** (conceptualized as sound quality or 'color,' also described further in chapter 3). Timbre helps us differentiate between the sounds of a flute and a clarinet, or to identify a voice on the telephone. A critical concept for understanding timbre is the 'overtone series' of a complex tone, which we turn to next.

Frequency relationships in the overtone series

The ***overtone series*** refers to the specific set of frequencies that are higher than the fundamental in a complex tone. The overtone series influences timbre in part based on the mathematical relationship among the frequencies in the overtone series. In many cases, the frequency of each overtone is an integer multiple of the fundamental frequency. For instance, in Figure 2.3 the first overtone is double the frequency of the fundamental and the second overtone is triple the frequency of the fundamental. This frequency pattern is similar to a harp string. When it is plucked, it does not just vibrate as a whole, but also in halves, thirds, quarters, fifths, and so on – each creating successively higher frequencies based on these integer ratios (2:1, 3:1, etc.). An overtone series like the one shown here defines a certain kind of complex tone, referred to as a ***harmonic complex tone***, and for these kinds of tones the overtones are often simply called ***harmonics***, with the term implying a certain mathematical relationship among the overtones. Harmonic complex tones are critically important for music because they convey a clear pitch, as we will discuss in chapter 3.

The numbering convention for frequency components varies depending on whether one adopts the more general distinction between fundamental and overtones (which can have any numerical relationship with each other), or that of a series of harmonics, including the fundamental (for the more restrictive class of harmonic complex tones). Figure 2.4 shows an example of the frequency components for C2 using music notation. There are 12 frequency components, starting with the fundamental frequency, which conveys the pitch at C2. By convention, the numbering of frequencies based on the fundamental and overtones starts at 0, so the fundamental is often referred to as F0. By contrast, numbering based on 'harmonics' (for a harmonic complex tone) starts at 1 (the fundamental being the first harmonic), so that the number of each component reflects its numerical relationship to the fundamental frequency (the second harmonic is double the fundamental, the next is triple, etc.).

Of course, not all complex tones are harmonic. As mentioned before, harmonic complex tones are characteristic of sounds with a pitch-like or 'melodious' timbre. On the other hand, sounds with a rough or 'noisy' timbre have overtones that are not so simply related to the

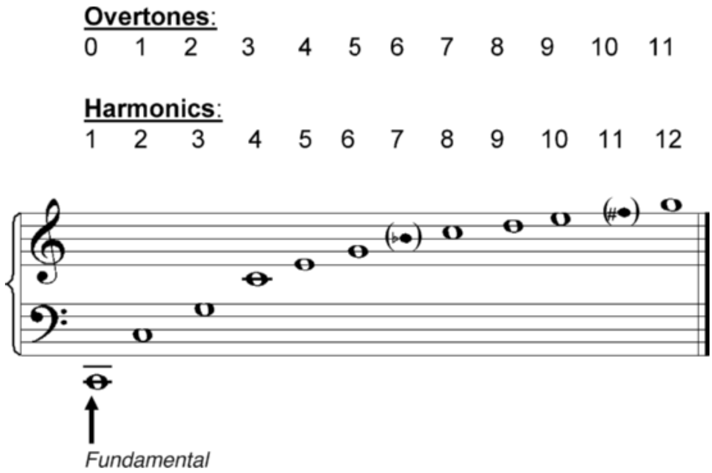


Figure 2.4 The harmonic or overtone series based on the fundamental C2, with harmonic and overtone numbers shown. Notes shown in parentheses denote approximate pitches. Numbering conventions for frequencies when considered as overtones or harmonics relative to the fundamental are both shown.

fundamental, called ***anharmonic complex tones***. Figure 2.5 shows an example of a complex tone (represented as in Figure 2.3) made up of sine waves that are not based on integer relationships. If played, the sound in Figure 2.5 would probably sound 'rough.' For instance, the static noise of a radio that is not tuned to a station comes from frequencies that are entirely unrelated to each other. Likewise, many percussion instruments (like drums and bells) include harmonics that do not conform to integer relationships.

In reality, the sounds we hear are almost never entirely harmonic, or entirely anharmonic (which would be white noise). However, this distinction provides a useful way to distinguish between different sorts of timbres. For instance, in the human voice, a smoother timbre (such as that of crooner Michael Bublé) is closer in structure to the kind of sound wave shown in Figure 2.3, whereas a rougher-sounding singing voice (perhaps Louis Armstrong or Tom Waits) would have more components that do not conform to such integer relationships.

Amplitude relationships in the overtone series

Differences among timbres are not only based on numerical relationships among sound frequencies. The second way that the overtone series determines timbre is through the amplitude associated with each frequency. Not all overtones necessarily emerge with the same intensity. If a harp string is plucked at its center, it is mainly the fundamental and the even-numbered overtones that will sound. Clarinet tones in the low register are almost completely lacking the first overtone. The sound each instrument produces is made up of a different composition of frequencies that in part give it a characteristic sound. This is another way in which Fourier analysis contributes to our understanding of sound. The bottom section of Figure 2.3 shows a representation of sound referred to as a ***power spectrum***, which plots the amount of energy (or power) associated with each frequency component in a tone (expressed as amplitude). Because energy is related to amplitude, the power spectrum can be thought of as a way of showing the amplitude of sine waves in a complex tone. The artificial

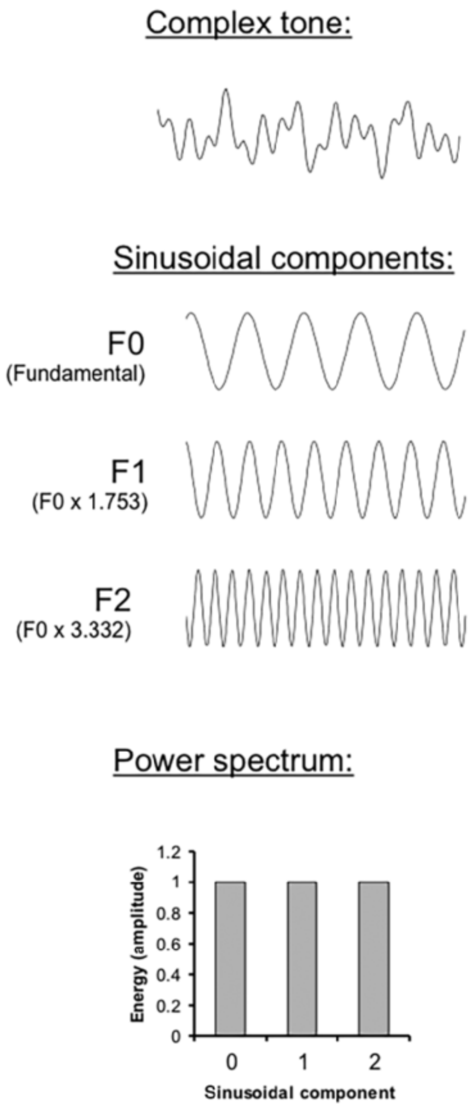


Figure 2.5 Waveform for an anharmonic complex tone, with its Fourier analysis below.

Source: Copyright © Peter Pfordresher.

tone in Figure 2.3 consists of components that are equally ‘powerful.’ What happens when components are not similarly powerful?

How the amplitudes of different partials are distributed defines an important measure of timbre: a sound’s ***spectral centroid*** (McAdams, 2013). The spectral centroid reflects the ‘center of gravity’ in the sound’s spectrum. Two examples are shown in Figure 2.6. When the spectrum is dominated by low-frequency energy, as in the left-hand sound wave in Figure 2.6, the centroid is based at a frequency close to the fundamental, and the subjective quality is usually considered to be ‘rich’ or ‘dull.’ By contrast, a higher centroid, as in the right side of Figure 2.6, can lead to sounds having a ‘bright’ or ‘nasal’ quality.

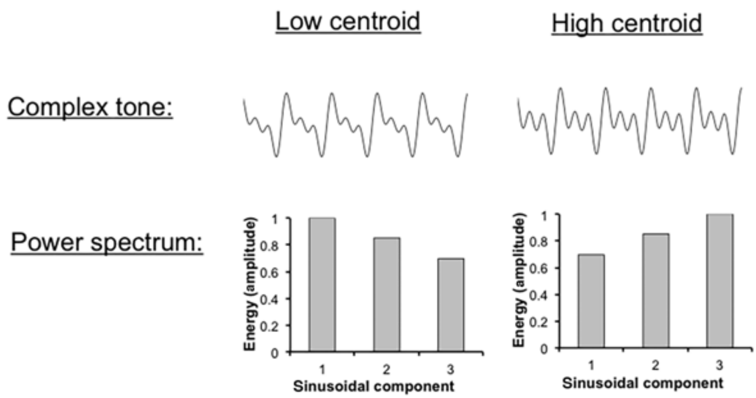


Figure 2.6 Waveforms for two complex tones differing in spectral centroid. The Fourier analysis is shown below each waveform.

Source: Copyright © Peter Pfordresher.

Thus far we have emphasized acoustical differences *across* instruments (and singers). However, it is important to note that the spectral structure of a single instrument is not constant, but can change as a function of pitch, loudness, duration, or expressive intentions.

An example of how pitch influences the power spectrum is shown in Figure 2.7. This figure was generated by recording a trombonist producing B♭ in three different octaves. The left side of the graph shows (complex) waveforms associated with each pitch, and the right side shows their power spectra (cf. Figures 2.3 and 2.5). Each B♭ is shown as F1 in the power spectra, along with its higher overtones. In principle, the power spectrum for each produced pitch could be identical because frequencies are plotted by harmonic number (1 for fundamental), so the fundamental frequency is always to the far left of the X-axis. However, this is not the case. The acoustic structure emphasizes relatively lower frequencies within the range of frequencies presented when the performer produces a higher pitch, and emphasizes relatively higher frequencies when the performer generates a lower pitch. These differences ultimately reflect the fact that a certain (intermediate) range of frequencies is likely to have a higher level of energy across many different produced pitches in the trombone. This range ends up being relatively high when one plays a low pitch and relatively low when one plays a high pitch. Such bands of high energy that are enhanced by the structure of a sound medium are known as ***resonances***, and may help the listener identify a specific instrument. We discuss the concept of resonance more in the next section.

Fourier analysis involves complex mathematics for which we have given a verbal sketch. Readers interested in this topic may wish to refer to Benson's (2006) *Music: A Mathematical Offering*, which includes a good chapter on Fourier analysis in a musical context, or Kammler's (2008) *A First Course in Fourier Analysis* (which, despite its title, is an in-depth exploration with applications to music and other areas).

The acoustics of musical instruments

Musical instruments (and we include the human voice among them) are devices for producing sound waves. Each species of instrument has its own acoustic characteristics, and the physical structures of instruments vary widely. However, almost all musical instruments depend on the

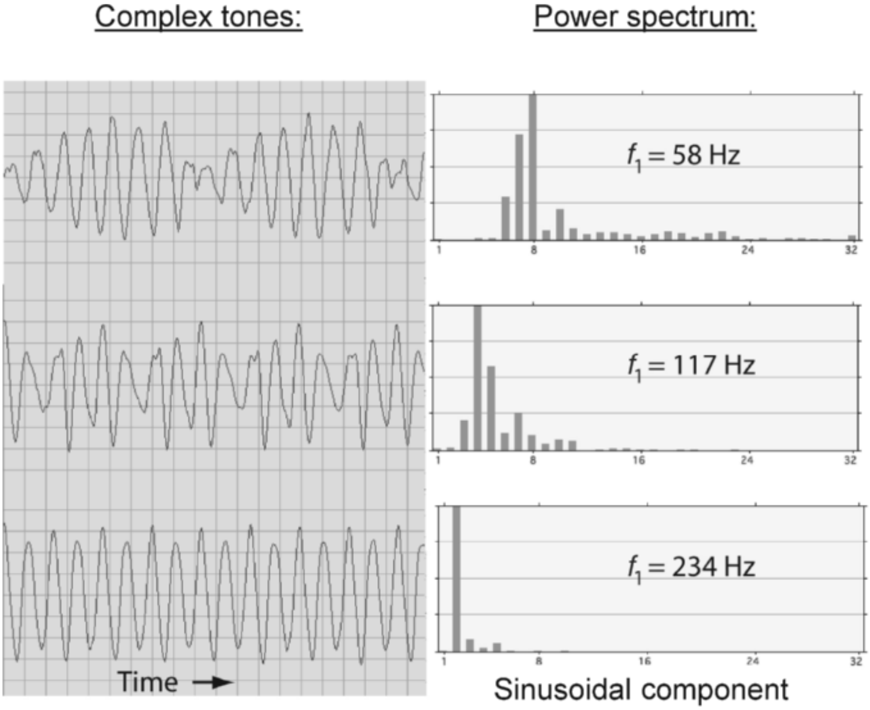


Figure 2.7 Waveforms (left) and power spectrum plots (right) for three tones produced by a trombone. (Power spectrum plots show frequency values as harmonic number rather than Hz.)

Source: Figure created by Matt Segars and composited by Rod Nave at the Department of Physics and Astronomy at Georgia State University, used with permission.

principle of ***coupled acoustics***. That is, in most musical instruments there are two vibrating devices: One *generates* the sound wave, and the other is ***coupled*** to it, and *amplifies* it.

The initial vibration that is activated by the performer – for instance, by blowing into an oboe or bowing a violin – is often of insufficient energy to be properly heard in a large space or at a distance. A ***resonator*** (such as a hollow tube, an air chamber, or a flexible soundboard) is 'coupled' to the basic device. The importance of coupling can easily be demonstrated by blowing the reed of a clarinet mouthpiece detached from the instrument, which will only produce a thin squeak. In the clarinet, the vibrations of the reed excite oscillations in the column of air in the body of the instrument. The air column is the resonator. In the double bass, the vibrations of the strings resonate in the hollow body of the instrument. The similarity between the terms ***resonator*** and ***resonance*** (introduced earlier) is no coincidence. Any given resonator is likely to serve as a better conduit for certain frequencies than others. These resonances, as mentioned before, help give a particular instrument or voice a characteristic timbre across many different pitches and loudness levels.

The groundwork for what we know about sound resonators and many other acoustic phenomena was laid by Hermann von Helmholtz in 1863, with his publication of *On the Sensations of Tone*. One of the leading physicists of the 19th century, Helmholtz's interest in music led him to propose a research paradigm for bringing the physical sciences into a working relationship with the arts. Realizing that the discoveries of physical acoustics (such as the ratios of modes of vibration of a taut string) were correlated with experiences of consonance, harmony, and other phenomena that were the basis of music, he looked for a bridging science. He found this in physiological acoustics, the anatomy and physiology of the ear, and neural links to the brain.

Most pertinent to the present discussion is Helmholtz's experimentation with sound resonators consisting of a large collection of hollow spheres (of many different sizes) made of glass or metal, with two tiny openings on each end. One opening was to be placed snugly inside the ear for an airtight fit, and the other was to be pointed at the source of a sound. Helmholtz discovered that if the sound contained an overtone with the same or similar frequency as the natural resonant frequency of the hollow space inside the resonator, the overtone would be amplified and would sound inside the resonator. Just as 'Helmholtz resonators' of different sizes and materials amplify different overtones, the properties of the resonators of musical instruments in part determine which overtones are amplified or suppressed. 'Brighter' tones occur when the higher overtones are present and amplified, and a 'duller' sound is produced when upper overtones are weak or missing.

Strings

We can group instruments in which sound is produced by setting a string into motion into three main groups. In one group there are 'bowed strings.' These are the instruments such as violins and double basses in which a string is continuously supplied with energy from a moving bow, and the vibration is maintained at the same pitch to give an enduring tone. In another group such as harpsichords, guitars, and mandolins, a string is set in vibration after brief contact with a plectrum or some other device that plucks the string (in the case of Jimi Hendrix, sometimes his teeth!). Then there is the third group of instruments in which the string is struck with a hammer, as in the piano, which also belongs to the family of percussion instruments. In most of these instruments, the hammer rebounds almost immediately off the string. Whether or not the string continues to vibrate depends on whether it is damped by contact with some material such as felt.

Not all vibrations of a string are controlled by the actions of the performer. One of the three strings of the long-necked string instrument the Indian *sitar* is not plucked, but vibrates in sympathy with the vibrations of the melody and drone strings. A subtler effect of *sympathetic vibration* occurs when an object vibrates not because it is stimulated by a direct action such as plucking a string, but because vibrations in a sound wave cause the object to vibrate due to the resonant frequencies of that object. The phenomenon of sympathetic vibrations can be demonstrated on a grand piano when playing one or more keys while pressing the sustain pedal (the rightmost pedal). The pedaling action removes the dampers from the strings, freeing many strings you are not striking with the piano hammer to vibrate in sympathy with the strings corresponding to the depressed keys.

Despite the variety of ways the strings are set in motion, all of these musical sounds follow the same principle of coupled acoustics. The vibrating strings radiate a weak sound initially, but the energy is transferred to the bridge and top plate, and into the body of the instrument and the back plate. In addition to other factors such as the thickness, length, mass, and tension of the strings, the particular characteristics of the resonating body play a critical role in the resulting sound's timbre.

Woodwinds and brass

The same coupling principle that lies behind the production of musical tones by stringed instruments also applies to wind instruments. A vibration produced in the mouthpiece sets a column of air resonating. It is this vibration that is transmitted to the listener by the air. There are several ways in which musical sounds are produced in woodwind instruments, a few examples of which are discussed next.

Edge tones are created by an action such as blowing across a hole and onto a sharp edge. The tones are produced because the upper and the lower planes of an edge produce asymmetrical vortices as the air stream passes over the edge. This asymmetry induces an oscillation in pressure, which is picked up and amplified by the adjacent air column in the body of the instrument. The effect only works if the incoming air stream is narrow, as for example from the pursed lips of the flute player (essentially a hole), or through the embouchure of the clarinetist (which is essentially a slit). This is another example of coupled acoustics, because the physical process at the edge is quite different acoustically from the wave transmission in the air column. But the instrument only produces musical tones because the two processes are coupled.

Reed tones are produced by the player setting one or two reeds vibrating, which in turn produces sound waves in the adjacent air columns. This is the system to be found in single-reed instruments such as clarinets and saxophones, and double-reed instruments such as oboes and bassoons. In the Scottish bagpipe, the player does not put the reed into the mouth, but blows into a reservoir in the form of a bag; when the bag is squeezed, sound is created as air passes through the reeds of three pipes creating drones and a fourth 'chanter' pipe providing the melody. Brass instruments use 'reed tones,' but the 'reeds' are the human lips; the vibrating motion that creates the sound is independent of the instrument, being freely controlled by muscle tension. Brass instrumentalists must expend a lot of effort developing the muscles of their lips as a way of controlling pitch height and intonation (good brass instrumentalists can play music simply by 'buzzing' their lips). Finally, we should notice that the human voice is a reed tone instrument, the vibrations of the vocal folds (also known as vocal chords) resonating in the air chambers of the respiratory system. The acoustics of the vocal resonators are complex; imagine the constantly changing shape of the resonating chambers of a vocalist singing a lyric such as 'Hallelujah!'

There are interesting variations in tone quality depending on whether the body of a wind instrument is of conical bore or even bore throughout, and how it flares at the end. A flared bell radiates the high overtones, which in part gives the trumpet its 'bright' sound. The length of the resonating columns of wind instruments also varies widely; generally, longer pipes produce lower or deeper tones. The tubing of a French horn when unrolled into a straight pipe, for instance, measures about 2.7 to 3.7 meters (depending on the type of horn). This is approximately two to three times the length of the pipe of a trumpet, accounting in part for the warm, rich tones it can produce. It is the length, shape, and materials of the pipe resonator that, to a large extent, account for the distinct sounds produced by different wind instruments.

Percussion

Unlike the violin or the clarinet, in which the performer continuously supplies energy to the system, percussion instruments receive energy in short bursts. A drumstick strikes the membrane of the drum in a single stroke. This stroke sets the relevant parts of the instrument vibrating at their natural frequencies, depending in part on the size and rigidity of the materials with which it is made. As the surface of the drum is flexible, the impact also sets into motion a wave disturbance that travels to other parts of the surface that were not directly struck. Thus a single strike of a drum generates many frequencies that bear no simple relationship to each other, producing an 'unpitched' sound. The induced vibrations then die away until the performer makes another stroke.

Drums, chimes, marimbas, and many percussion instruments have resonating chambers or tubes. The coupled acoustics of the *saron*, one of the pitched percussion instruments in the Indonesian *gamelan* orchestra, can easily be deconstructed as it consists of a few loose metal keys, simply placed over a shallow wooden box, which serves as a trough resonator. Without it, the *saron* would barely be heard over the full-percussion *gamelan* that often performs in outdoor venues. The piano's main resonator is the soundboard; essentially, it repeats the vibratory motions of the piano strings and must be carefully crafted if the instrument is to produce a rich sound. However, not all percussion instruments have resonators. For instance, vibrating plates such as cymbals and gongs have no built-in 'amplifiers.'

Across the huge diversity of musical instruments which human ingenuity has created, the basic principles are the same. With very few exceptions, instruments produce pitched sound waves, from which the listener extracts the basic musical qualities of pitch, loudness, timbre, and variations in duration (rhythm). Almost all instruments depend on the principle of coupled acoustics to generate sound waves that are clearly audible at some distance from the performer, and the various resonators contribute their characteristic properties to the sound wave that ultimately emerges. However, the quality of sounds produced by the instruments and voices is also affected by the characteristics of the physical environments into which they are released. In the final section in this chapter, we consider the sound fields in which music is commonly performed and appreciated: concert halls, opera houses, and other performance venues.

The acoustics of musical venues

In 1895, Wallace Clement Sabine, professor of mathematics and philosophy at Harvard University, was summoned to address a practical problem: Listeners were having a hard time understanding what speakers were saying in a lecture hall in Harvard's Fogg Art Museum. After performing a series of tests, Sabine determined that sounds in the hall were sustained within an audible range for a long duration before decaying, making speech indistinct. Sound-absorbent materials were placed in the room, and the problem was remedied. Prior to this time, buildings were not usually constructed with acoustics in mind. The first music hall to be designed with the help of acoustic engineering was the New Boston Hall (Boston Symphony Hall) in 1900, for which Sabine applied formulas he developed for calculating ideal reverberation times. He was soon regularly consulted to assist in building design and address acoustic problems of completed structures. Sabine's work laid the foundations for ***architectural acoustics*** (a field that encompasses the analysis, design, and control of sound in a building or other structure), which was later continued in the realm of music performance venues by Beranek (2004, 2007).

Today, the study of the acoustics of a concert hall or opera house has come to encompass more than just the (rather complex) physics of sound propagation in an enclosed space. It has come to include the consideration of features that affect the perception of music, and often also the subjective evaluations of the listeners and performers. The best conditions for hearing an orator can be fairly well defined in acoustic terms, as the focus is often on the clarity and intelligibility of speech. However, a host of other factors, such as the type of music being performed and even the reputation of the music hall, are relevant to how music sounds to a particular listener or performer. Following important work by Yoichi Ando (1985) and others on the *subjectivity* of architectural acoustics, much of the research on the acoustics of music performance venues takes listeners' and performers' subjective preferences into account (e.g., see Gade, 2015). Although not the focus of this chapter, there are also many social variables (such as the degree to which patrons feel socially comfortable in the venue and familiar with the orchestra), which cannot be isolated from the perceived musical quality of the concert experience (Pitts, Dobson, Gee, & Spencer, 2013).

Studies addressing the quality of the experience of music under different conditions can be carried out in several ways. It is possible to create artificial sound fields with properties that mimic the structures of the sound fields one would find in a concert hall, church, or other enclosed space. For instance, different materials can be placed inside an ***anechoic chamber*** (an insulated room designed to completely absorb reflections of sound and thus be free of echoes) to shape the way sound fields are created, and listeners may be asked to judge the quality of the sound piped in through loudspeakers. Another method is to create ***binaural musical recordings*** (by employing a model of a human head with stereo microphones inserted in each ear, in order to record sound traveling to both ears) to quite faithfully reproduce musical performances in different halls. Listeners can then compare these recordings inside an anechoic chamber.

Increasingly, researchers are also using ***computer modeling*** (involving computer programs to simulate real-world situations and processes) to simulate how sound is propagated in a room or concert hall. For instance, the LIVE (large interactive virtual environment) lab at McMaster University uses computer modeling to create a virtual acoustic environment that can be controlled by researchers to simulate how a piece of music sounds in different performance spaces (such as a jazz club, subway station, or cathedral), while measuring heart rate, breathing rate, brain activity (by electroencephalography), and many other responses of the performers and audience.

None of these procedures, however, captures the complete experience of sitting in the performance hall during a live concert! Therefore, an alternate method that has been used is to ask listeners to give their impressions of the acoustical qualities of live performances in music performance halls of which the sound fields and acoustic parameters have been studied. Leo L. Beranek (2004) used this approach in an extensive study of 50 of the major concert halls and opera houses of the world, originally reported in 1962 and subsequently revised in an expanded volume in 2004. He set about interviewing conductors, music critics, experienced listeners, and performers about their experience of music in various venues. Then, using standard acoustical concepts and techniques, he also measured the physical attributes of the halls in order to also collect more objective data. As we shall see later in this chapter, he found some consistency in listeners' and performers' preferences in some of the world's great music halls.

Direct and reflected sound

Our discussion of the acoustics of venues for musical performance focuses on a few basic concepts as they apply to the quality of listeners' and performers' experience in music performance venues: directed and reflected sound, sound absorption, and reverberation time. Throughout, both physical (or objective) and subjective parameters will be considered.

Direct sound travels directly from the source to the listener; it contains auditory information in an uncontaminated form. The clarity of the sound of an orchestra depends heavily on direct sound. ***Reflected sound*** reaches the listener by 'bouncing off' one or more surfaces such as walls, ceilings, pillars, and sound baffles. Depending on the time interval between the arrival of direct and reflected sound at the ears of the listener, the added reflected sound may add a pleasant richness to the musical tones – or it may 'muddy' up the sound and even create distracting echoes. At high reflection levels, the delay associated with these echoes is around 50 milliseconds (50/1000ths of a second; Gade, 2007). This time delay may seem small, but for complex musical passages it can interfere considerably with the experience of music.

An important measure is the ***initial time delay gap (ITDG)***, defined as 'the time at which the first reflection is heard after the direct sound' (Beranek, 2007, p. 4). An ITDG of less than 25 milliseconds has been found in the best concert halls, including the Concertgebouw in Amsterdam and Boston Symphony Hall. This allows listeners to detect clear onsets of musical tones, to hear successive tones distinctly, and to localize where sounds are coming from on the stage. If the ITDG is greater than about 35 milliseconds, 'the hall will sound like an arena, with a lack of intimacy' (p. 4).

There is also the clarity of sound to consider, or the extent to which musical tones or a singer's lyrics sound clear and distinct. The ***clarity index*** is a measure of the ratio of early sound energy (from ***early reflections***, arriving within about the first 80 milliseconds of the direct sound) to late sound energy (***late reflections*** typically arriving after 80 milliseconds). The preferred clarity index of a performance venue varies with the characteristics of the music being played; contrapuntal Baroque music demands higher clarity than music of the romantic period (Reichardt, Alim, & Schmidt, 1974, p. 243) so that the interlaced lines or 'voices' can each be heard distinctly.

Two-thirds of the world's concert halls that were most highly rated by conductors, music critics, and experienced concert-goers in Beranek's (2004) study are 'shoebox'-shaped. The narrow width of the halls allows for stronger ***lateral reflections*** (from side to side) to complement the sounds coming from directly in front of the listener (from the stage). Strong lateral reflections tend to produce a sensation of being 'bathed' in the sound, and have been shown to enhance musical dynamics (i.e., expressive variations between soft and loud) for listeners (Ando, 1985; Pätynen, Tervo, Robinson, & Lokki, 2014). If a shoebox-shaped hall is too wide (more than 25 meters across), the sound produced will tend to be muddy. However, if the hall is too narrow (less than 15 meters wide), the direct sound and first reflected sound may merge so that the sound reflected from the closest wall could mask and diminish the vibrance of the sound coming directly from the stage (Beranek, 2015).

The 'shoebox' design should not imply a long narrow hall with flat walls and plain ceilings. Sound is dispersed more uniformly throughout a space when there are some irregularities in the 'box,' such as coffered (ornamental, recessed) ceilings, niches in the walls, balconies, statues, and textured surfaces that can diffuse or diffract sound waves. Irregularities in the lower part of shoebox-shaped halls diffuse reflected sound and create the impression of 'warmer' tones, and those in the upper part of halls enrich the reverberant sound (Beranek, 2015).

Fan-shaped halls are often more problematic. The fan shape leads to progressively widening walls toward the back of the hall, which directs lateral reflections away from the listeners. This may weaken the sense of ***acoustic intimacy*** (the degree to which sounds seem to be coming from nearby rather than remote surfaces), ***listener envelopment*** (the sense of being surrounded by reverberant sound coming from all directions), and ***warmth*** (provided by transmission of low frequencies such as bass tones). Further, while the first reflections arrive from the narrow parallel walls in shoebox-shaped halls, the first reflections from halls that fan out from the stage are likely to come from the more distant path of the tall ceilings – thus making for a longer reverberation time. Concave walls behind the audience also tend to send reflected waves back to a point on the stage, and can create echoes – an annoyance to many performers (especially instrumentalists).

To give an interesting case study, a notable exception is the fan-shaped *Aula Magna* at the University of Caracas in Venezuela, which receives good reviews from performers and critics. To address the potential acoustical problems stemming from the nonparallel walls and curved features of the architectural design, it was decided during construction that sound-reflecting panels often referred to as ‘clouds’ (covering an area equivalent to 70% of the ceiling) were to be hung below the ceiling and along the side walls. Rather than using standard rectangular panels, sculptor Alexander Calder was commissioned to create suspended panels in the dynamic abstract shapes characteristic of the geometric-shaped mobiles for which he is known. The result, shown in Figure 2.8, is both acoustically and visually pleasing!

However, pure laws of physics alone do not always predict audience preferences. Eight of the nine top-rated opera houses in Beranek’s (2004) study are horseshoe-shaped. He noted that the horseshoe shape is not acoustically ideal (the curved wall behind the audience directs sound back to the center of the stage, and balconies create areas of low resonance underneath them), and yet seemed to be preferred by listeners and performers for other qualities. For example, this arrangement brings performers and audience closer together, which increases ***visual clarity*** so that facial expressions and gestures can be clearly seen, and ***acoustical clarity*** for the intelligibility of the lyrics and speech. The slight echo common in halls of this shape provides feedback to singers, and does not seem to distract listeners.

Sound absorption

Not all sound waves reach the ear directly (direct sound) or bounce back from dense surfaces (reflected sound) in a contained space. Some become trapped in materials such as ceilings

Figure 2.8 The interior of the Aula Magna at the University of Caracas in Venezuela, featuring ‘clouds’ by sculptor Alexander Calder.

Source: Published courtesy of Leo Beranek, *Concert Halls and Opera Houses* (Springer, 2004).

and walls, stage curtains, carpeting, and seat upholstery – and reflect little energy back. This sound has been *absorbed*.

Absorption materials commonly found in music performance venues can be classified into two kinds. ***Porous absorbers*** (curtains, theater seats, carpets) absorb high-frequency sounds more efficiently than bass sounds. ***Resonant absorbers*** (such as wood panels) are set into vibration by the energy released by the sound source and respond to low-frequency sounds. By conducting numerous tests of the sound absorption of different materials, Sabine determined that overall ***absorbing power***, the key to reverberation times, varies not only with the material reflecting sound, but with the frequency (pitch) of the instrument. For example, the absorbing power of a given surface for the higher register of the violin is nearly double that of the same surface for the lower register of the double bass (Sabine, 1922).

Audiences also absorb sound; imagine the heavy fabrics of formally attired concert audiences of the 1700s and 1800s! Audience absorption is a tricky variable, as one cannot always predict the number of occupants who will attend a performance, nor the exact arrangement of occupants filling the available seats. The audience area is larger if the audience is seated on a slope, and thus absorption is affected not only by the size of the crowd, but by the dimensions and plane of the seating space. The seats themselves also absorb a lot of sound, which is one reason why back rests in most concert hall seats are usually low (ending below the shoulder) and chairs are not as plush as seats in a movie theater. Chairs with lightly cushioned seats and wooden arm rests reflect more sound and absorb less high-frequency sounds, so the delicate tones of chamber music are more intense and brilliant. On the other hand, heavily upholstered chairs can keep the music from sounding too loud in small performance venues by absorbing more sound (Beranek, 2015).

Reverberation time

Sabine was a pioneer in identifying reverberation time as a critical factor in indoor acoustics. ***Sabine's formula*** (i.e., his original equation for calculating reverberation time, *RT*) is written as:

$$RT = 0.161\mathrm{V/A}$$

where V is the volume of the room in cubic meters, and A is the total absorbing power in sabins. Sabine's measure of reverberation time was the time it takes a sound to decay until it is barely audible, as sensitive methods to determine sound levels had not yet been devised. He made painstaking systematic measurements of the time it took sound to decay, starting from 1,000,000 times the first audible sound level to silence (Sabine, 1922, pp. 60–68). Today, ***reverberation time (RT)*** is defined as the length of time for a sound to decay by 60 decibels. Although alternate measures such as ***early decay time*** (***EDT***, based only on the initial part of the decay) have since been devised, RT is still regarded as one of the most important basic acoustic parameters.

Performance halls range from reverberant to dry – that is, from those in which there are many complex reflections to those with absorbent walls and ceilings, which reduce the loudness and longevity of these reflections. From the perspective of performers, two important concerns are ***ease of ensemble*** (the degree to which performers can hear themselves and others playing together) and ***support*** (the degree to which the room facilitates the musicians' efforts to create tones and fill up the space). Early sound reflections in the stage area are critical for ease of ensemble, while both early and late reflections are important for good support. When performing in a ‘dry’ space, it is difficult for instrumentalists and vocalists to feel like they are ‘filling up’ the room with music, as the sound disappears quickly. They may then sing or play with more force, and the tone quality may be compromised (Gade, 2015).

Instrumental and vocal ensembles must also give more precise performances in a room with a short reverberation time, as asynchronies are more discernible. A choir singing lyrics such as ‘to tell a tale of tragedy’ must really stay together to sound crisp in a dry space! On the other hand, long reverberation times in large spaces may make it difficult for performers to hear themselves and each other, as ***horizontal clarity*** (the distinctiveness of tones played successively) and ***vertical clarity*** (the distinctiveness of tones played simultaneously) are diminished. Indeed, long reverberation times have been found to affect the tempo of performance for choral singers, as they tend to sing more slowly and with less precise timing in reverberant rooms, although intonation (pitch accuracy) may not be affected (Fischinger, Frieler, & Louhivuori, 2015).

The highest-rated concert halls in Beranek’s (2004, 2007) studies have a reverberation time of 1.8 to 2.0 seconds, which seems ideal for orchestral music, while a shorter RT of 1.24 to 1.6 seconds seems to be ideal for the best opera houses. Venues designed for speech (such as oration, and plays) generally require the shortest RTs of 0.7 to <1.0 second, as it is critical for the audience to perceive rapidly unfolding speech sounds. In listening to music, however, the *loss* of some sounds (the rasping of the cello bow or the hiss of air expelled from a clarinet) due to longer reverberation times may actually enhance the listener’s appreciation of the music. Further, some degree of blending together of the sounds of an orchestra or choir is essential to a music performance. Auditoriums with modifiable size and flexible reverberation times (e.g., with chambers that can be opened or sealed, or removable panels) are practical designs for performance venues that must accommodate speech, as well as vocal and instrumental performances.

Optimal reverberation times vary for different instruments and vocal ranges, and musical repertoire. Even within the same ‘family’ of instruments, ideal reverberation times differ somewhat; pianists prefer halls with shorter reverberation times, whereas pipe organists prefer more reverberant spaces (Veneklasen, 1975). La Scala in Milan has an RT of 1.2 seconds, while Wagner’s Festspielhaus in Bayreuth has a much longer RT of 2.2 seconds. It is not hard to see why this difference should emerge, considering the characteristics of Italian and German opera! One might compare the lucidity of Verdi’s operas and his focus on the singers and text with the lush, expansive orchestration of Wagner’s operatic works.

Very large auditoriums are rarely effective for music performance. In his study of architectural structures for the performance of music from the 17th century to the present, Forsyth (1985) notes that the Royal Albert Hall in London (built in 1871) was *10 times* larger in volume than most concert halls of its time, and attributes many acoustical problems it encountered through the years to its immense proportions. In very large modern spaces such as indoor arenas which may house audiences as large as 10,000 or 15,000 or more for ‘pop’ and rock concerts, performers must rely on electronic amplifiers and loudspeakers to carry and disperse sound, and low-frequency reverberation is particularly difficult to control. Generally, low-frequency absorbers are not as efficient as those that absorb high frequencies, which is why audiences in such venues often find themselves in a wash of persistent low-frequency sounds.

Bidirectional influence: Music and architecture

Our discussion has focused on how considerations about music performance have shaped the way we design and construct buildings, but the effects may have been *bidirectional.*

Sabine once argued that the acoustics of a sound space are so critical to effective performances that the architectural traditions of different eras may have fundamentally shaped the development of music (cited in Forsyth, 1985). The cavernous, highly resonant stone buildings of the Romanesque period allowed vocal tones to linger, supporting the exploration of rich vocal harmonies characteristic of choral music of that time. As the classical outdoor amphitheater evolved into the roofed horseshoe-shaped concert building, the improved horizontal clarity may have lent itself to the development of ornate contrapuntal Baroque music with its complex interplay of melodies (Forsyth, 1985). An analysis of symphonic scores suggests that some classical composers took the effects of reverberation in performance venues into consideration in their composition and orchestration (Meyer, 2015). It is intriguing to consider the possible *mutual influence* of the construction of musical buildings and the construction of musical works.

Coda

At first glance, the science of acoustics does not seem very daunting. Wave trains in the air and the overtones of vibrating strings are not difficult concepts to master. Acousticians have succeeded in extracting many listener-relevant properties of such sound fields, so that we can form fairly clear ideas of how musical experience is related to the physical properties of the field. The complexity of sound fields is directly related to the structure and properties of the bounded spaces in which music is played and heard. But here, mathematical analysis and ascertainable principles begin to part company. Using the broadly defined acoustical concepts described in this chapter, it is possible to draw some correspondences between architecture and audience. However, one of the lessons to be learned from the research on the acoustics of performance venues is how difficult it has turned out to be to design a 'perfect' acoustic environment from the principles of the science of acoustics! There is much about the experience of performing and listening to music that cannot be captured by physical laws alone. It is to the psychological and more subjective qualities of musical sound, and the pathway from the source of a sound to the ear and auditory cortex, that we turn our focus in our next chapter.

3 Auditory perception and the neurophysiology of hearing

Passport photographs for many countries require both ears to be clearly visible. This is because the ears, though not quite as unique as fingerprints or snowflakes, are distinctive features of a person's appearance. If you were to gather a dozen of your friends and peer closely at their ears, you would be likely to find great variability in the shape, size, and particularly in the convolutions or patterns of ridges of their ears! In fact, outer ears are so different with respect to their shapes and irregularities that if your outer ears were switched with someone else's, you would probably have trouble accurately identifying exactly where sounds were coming from in the environment around you.

This chapter focuses on the workings of the marvelous ear, to reveal the mechanisms in the auditory system that enable people to hear the musically salient aspects of sound. First, we begin with an examination of pitch, loudness, duration, and timbre as the four dominant perceptual properties of the heard sounds from which music emerges as an auditory experience. In the second section, we describe the main structures of the ear and the pathways leading from the ear to the brain. How do the ears and the relevant parts of the nervous system extract these features of sound? At the end of the chapter, we conclude with a brief discussion on the topic of cochlear implants and implications for the perception of music.

Perceptual qualities of sound

Chapter 2 described how a sound wave is propagated, and explained that sound waves can be characterized by their ***physical*** properties: *frequency*, *amplitude*, and the *power spectrum*. As discussed briefly in that chapter, these physical properties of a sound wave correspond to the ***perceptual*** (i.e., subjective) qualities: *pitch*, *loudness*, and *timbre* respectively, as shown in Table 3.1. In addition, there is also the *duration* of the acoustic or auditory event, as sounds must unfold in time.

The distinction between physical stimuli and our perceptual experience of those stimuli is an important one in psychology. Our brains do not reconstruct all the physical properties of incoming stimuli. If they did, we would see X-rays, and hear radio waves, and the result

Table 3.1 Physical and perceptual properties of a tone.

Physical properties of sound waves	*Perceptual properties of tone*
Frequency (hertz)	Pitch
Amplitude (decibels)	Loudness
Power spectrum	Timbre

would likely be confusion. Instead, our brains have adapted to process those aspects of the physical world that benefit our survival, and they do so in a way that highlights those features that matter. In this chapter, we will discuss some basic ways that auditory perception involves a transformation of the physical sound wave described in chapter 2. The chapters that follow further explore how our brains reframe the experience of sound in a musical context.

Pitch

The perceptual experience of ***pitch*** is related to the ***frequency*** of vibrations in sounds. The ability to perceive pitch thus relies on the ear's ability to encode frequencies from physical stimuli. The pitches that most human beings can detect range from about 20 to 20,000 Hz (based on the unit ***hertz***, which refers to *the frequency of a wave expressed in cycles or oscillations per second*). This capacity varies widely from person to person and declines with age. Starting earlier, though often not really noticeable until about age 60, ***presbycusis*** (i.e., hearing loss associated with aging) results in a loss of sensitivity especially to high-frequency sounds.

While the range of intact human hearing is about 20 to 20,000 Hz, most musical pitches fall within the range of approximately *20 to 4000 Hz*. For instance, the fundamental frequencies of a grand piano with 88 keys range from 27.5 Hz to 4186 Hz. Towards the extreme ends of the spectrum of musical tones, the lowest pipe in a pipe organ is about 16 Hz (and is often felt more as a rumble or vibration than a tone), and the highest note of the piccolo may reach around 4500 Hz, producing a shrill sound that is uncomfortable to some listeners. However, the threshold of human sensitivity to pitch extends far beyond the piccolo's top register, enabling us to also appreciate some of the high overtones in its timbre. To consider the highest pitch that a college student with intact hearing can detect (20,000 Hz), imagine a grand piano extended by a little over two more octaves of keys. Sadly, such pristine hearing among young people is not common these days, a point we will return to later.

In the previous chapter, we discussed how most musical sounds come from a complex pattern of vibration that can be broken down into many component frequencies. Even a single musical tone, created by depressing one key on a piano, includes many different frequencies because of the many mathematically related ***overtones*** (or ***harmonics***) that also sound (at different relative loudness for different instruments, as described in chapter 2). Yet the perceived quality is of one single pitch. Relatedly, one of the mysteries of auditory perception, directly relevant to music, is that we typically hear one tone as having one pitch, but can still make out distinct pitches when multiple tones form a cluster, for instance when a pianist plays a chord. Much research has been devoted to understanding how the auditory system extracts a single pitch from a complex combination of frequencies – the reverse of Fourier analysis (described in chapter 2) – leading to theories of pitch processing in the auditory system. Such theories have largely been motivated by the physiology of the cochlea. We will therefore discuss them later in this chapter, accompanying a description of how the cochlea transduces the sound wave.

The simplest account of pitch processing, which works in many but not all cases, is to say that perceived pitch is linked to the ***fundamental*** (i.e., the lowest frequency in a complex tone), while the higher overtones contribute to the timbre of that tone (see chapter 2). However, this generalization does not always apply. Consider listening to someone singing a song on the telephone. The frequency range of most telephones is limited, and as such does not go low enough for many fundamental frequencies of low voices. Yet a listener will have the impression of hearing a voice at about the same pitch (though not the same timbre) as if the singer were physically present. This experience is connected with a curious effect, the phenomenon of ***residue pitch*** or ***virtual pitch***: that is, under some conditions, a missing fundamental will seem to be heard solely on the basis of hearing its overtones (Schouten, 1940; Terhardt, 1974). The same phenomenon is at play when listening to a pocket-size transistor radio. The fundamental frequencies of the lower tones are often not audible, and yet the proper pitches are usually heard as the fundamental is 'inferred' from the overtones.

Consonance and dissonance

Another important concept related to pitch perception, motivated by physiology but grounded in behavioral studies, is the ***critical bandwidth*** (or ***band***). The critical band refers to a range of frequencies that evoke a similar response in the auditory system, and bears on the degree to which our auditory system responds selectively to different frequencies. Frequency is a continuous variable, with an infinite number of possible values, and so it would be inefficient for our auditory system to respond selectively to every possible value. So instead, it seems that our auditory system responds similarly (e.g., with a similar cochlear response) to frequencies in close proximity to each other, and frequencies that evoke a similar response are said to fit within the same critical band. In some cases, this 'shortcut' of the auditory system can influence our perception of music. The most prominent example is in the case of musical dissonance.

The terms ***consonance*** and ***dissonance*** refer to a basic and important subjective continuum in music – reflecting the degree to which tones that are played simultaneously (called a *harmonic interval*) sound pleasing and relaxed, or displeasing (Helmholtz, 1863). In a highly influential paper, Plomp and Levelt (1965) demonstrated that the degree of perceived dissonance for intervals formed by complex tones may come from interactions among the upper overtones. Specifically, they found that perceived dissonance is maximal when there are overtones that fit *within the same critical band*, while not being identical in frequency. The greatest dissonance is associated with approximately 25% of the bandwidth, which happens for highly dissonant intervals like the tritone (e.g., C and F#). In contrast, the overtones for highly consonant intervals (e.g., C and G, a perfect fifth) are either identical in frequency or do not fit within the same critical band.

Thus, one of the key insights imparted by Plomp and Levelt's foundational work is that in order to understand consonance and dissonance, we cannot merely consider the relationship between the two fundamental frequencies that make up an interval – but must look further to the relationships between their overtones. When the fundamentals of the tones, or their constituent overtones, are so close together that they fit within the same critical band, they pose a challenge for the auditory system to reconcile; this tends to produce a sense of roughness that we may perceive as 'dissonance.'

Although Plomp and Levelt's contribution offers a good starting point for the understanding of how we experience consonance and dissonance, more recent work suggests familiarity may play a role that is at least as strong – or perhaps even stronger – than the role of critical bands. Whereas the critical band is a common feature of all human auditory systems, McDermott, Lehr, and Oxenham (2010) found large individual differences in the degree to which this construct predicts the experience of consonance across individuals. Perhaps most telling was the fact that musically trained individuals perceived consonance or dissonance based on whether two pitch categories – and not necessarily their overtones – are commonly considered to be consonant in musical practice. That is, a musician tends to hear a C-major chord (consisting of C, E, and G) as 'consonant' even if the component frequencies of each note are re-organized to create roughness. By contrast, other listeners (mainly nonmusicians) were more strongly influenced by the composition of overtones, and would not always hear C major as 'consonant.'

Such findings introduce the possibility that familiarity with musical rule systems influences our perception of consonance, perhaps even more strongly than purely physical factors (McLachlan, Marco, Light, & Wilson, 2013). More exciting new work on crosscultural differences in consonance perception is discussed in chapter 15, again suggesting that our knowledge adds a lot to the basic processing discovered by Plomp and Levelt.

Loudness

It may be surprising to learn just how sensitive the auditory system is to differences in the amount of energy that is transferred by a sound wave (i.e., in the ***amplitude*** of the physical sound wave). In fact, if the ear were more sensitive than it is, we would actually perceive the collisions among air molecules! How loud a sound is perceived to be matters a good deal in music, as in other contexts in which what we hear is of significance. The more energetic the pressure wave – that is, the greater its amplitude – the louder the sound heard will be as the wave impacts on the ear.

The perceived loudness of a sound is based on its ***intensity***. The most common way to represent different levels of intensity is through the decibel scale, which is based on a formal mathematical definition. Specifically: *Intensity in decibels (dB) of a sound is defined as 10 × the logarithm of ratio of the energy of that sound to a measure of static sound pressure*. Thus, if we give sound pressure (usually estimated as having an energy of 10^{-12} watts per centimeter squared) a value of 1, the decibel measure is 10 times log 1/1, that is 10 × log 1. Log 1 is 0 dB. The faintest discernible sounds approach this measure. Suppose we double the energy of the impinging sound. The measure will now be 10 × log 2/1, that is log 2. Log 2 is 0.3, so the decibel measure is 3 dB. If we double the energy again, leading to a measure of 10 × log 4/1, we get 6 dB, and 10 × log 8/1 gives 9 dB.

Although the decibel scale better approximates perceived loudness than would the raw energy ratio (Moore, 2012), it is important to note that the decibel scale measures the *physical signal*, not a *perceptual experience*. One's experience of loudness is only roughly proportional to the energy of the impinging sound wave (e.g., perceived loudness approximately doubles for every 10 dB, but only for moderate intensities; Hellman, 1976). In some instances, the physical measure of intensity in decibels does not correlate perfectly with perceived loudness in real-world listening conditions. This is because how loud a sound is perceived to be depends not only on the amplitude of the wave, but also on its frequency. For this reason, ***A-weighted decibels (dBA)*** are typically used to express relative loudness as perceived by the human ear; these are based on weighting curves that accommodate for how the ear is less sensitive to sounds at very high and very low frequencies, for instance. (For a discussion of other loudness scales such as *phon* and *sone*, see Rasch & Plomp [1999].)

To give some reference to 'real-world' sounds in decibels (hereafter assumed to be A-weighted): A whisper at a distance of 3 feet away is about 10 to 15 dB, the ambient noise inside a restaurant is about 45 to 55 dB, conversation at a distance of 3 feet is approximately 65 dB, and a hammer hitting a steel plate at about 2 feet away is 114 dB (Pierce, 1992, pp. 119–120). The human auditory apparatus allows people to discriminate loudness differences of approximately 1 dB for frequencies between 200 and 4000 Hz (Yost, 2000). Bats, dogs, and many other animals have much more sensitive auditory systems.

Application: Music-induced hearing loss

Many countries set regulations for noise exposure in the workplace. For instance, the Occupational Safety and Health Administration (OSHA, 1983) in the US recommends a

permissible exposure limit (PEL) of no more than 90 dB over an eight-hour period for working environments, as prolonged exposure to loud noise has been shown to lead to hearing loss. Orchestral performances have been measured at sound levels of up to 112 dB (Sataloff, 1991), and vocal performances during opera can reach 105 dB for voices alone without accompanying instruments (Laitinen, Toppila, Olkinuora, & Kuisma, 2003). Even the average sound level of a single instrument or solo voice music produced by a music student in a small practice room has been recorded at 87 to 95 dB (Phillips & Mace, 2008). Thus, it is not surprising that some studies have found the incidence of hearing loss to be greater among musicians than nonmusicians. For instance, Axelsson and Lindgren (1981) found that 42% of adult orchestral musicians exhibit greater hearing loss than expected for their ages. Further, 45% of undergraduate music students have detectable hearing loss, suggesting that the damage may already be manifested during the college years (Phillips, Henrich, & Mace, 2010).

The phrase ***music-induced hearing loss (MIHL)*** refers to loss primarily caused by engagement in playing music, listening to music, or to exposure to music in the environment such as in a club. As it is impossible to separate this from exposure to other sounds in musicians' lives, it is also more broadly referred to as a subset of ***noise-induced hearing loss (NIHL)***. MIHL may be instrument-specific, as some studies have found greater hearing loss in the left ear for violinists and in the right ear for flute players (e.g., Axelsson & Lindgren, 1981; Ostri, Eller, Dahlin, & Skylv, 1989) though other studies have not shown clear patterns corresponding to instruments (e.g., Phillips et al., 2010). Those seated in front of loud sections of the orchestra (especially brass or percussion) may also be at greater risk for developing MIHL. Rock and pop musicians often suffer more severe damage in the ear closest to the source of sound, such as a loudspeaker or amplifier (Sataloff, 1991).

At high intensity, even brief exposure to very loud music can lead to permanent damage to one's hearing. For instance, Park (2003) found maximum noise levels in *karaoke* singing environments in Korea to exceed 95 to 115 dB, and found up to 8 dB of significant hearing loss (centered at around 4000 Hz) in listeners after less than two hours of exposure! Unlike hearing loss due to aging (presbycusis), which tends to affect sensitivity to higher frequencies, hearing loss due to exposure to loud music affects sensitivity to music, as it begins much closer to the region of frequencies within the musical range. Typically, music-induced hearing loss is found at around 4000 Hz or 6000 Hz (Phillips et al., 2010). As musicians pride themselves on their ability to hear tonal nuances, this presents a significant loss to both production and reception of music. For instance, even tones in mid-register may lose their fullness or richness, as the high overtones would be reduced or lost.

A study showed that early noise exposure in mice produces pathological changes in the auditory system that make the inner ear more susceptible to the effects of aging (Kujawa & Liberman, 2006). If this also holds for humans, this finding suggests that early exposure to loud noise (including music) could predispose a person for more deleterious effects of presbycusis later in life.

Personal listening devices and MIHL

A growing body of research focuses on the risks involved in the widespread use of portable listening devices (PLDs) such as MP3 players and iPods, due to their high output levels and loud listening habits reported by users. For instance, about one-third of college students reported occasionally listening to MP3 players at 100% listening level (Hoover & Krishnamurti, 2010), which has been measured at over 125 dB on some PLDs (Breinbauer et al., 2012).

PLD users are at greater risk for developing irreversible hearing loss, especially when using devices in environments in which the background noise is louder than 65 dB (Jiang, Zhao, Guderly, & Manchaiah, 2016), because listeners tend to turn the music up to compensate for the ambient noise. For instance, commuters in the New York City transit system listened to their PLDs at a maximum of 106 dB when standing on subway platforms and 112 dB while riding inside subway cars, in order to hear the music (Gershon, Neitzel, Barrera, & Akram, 2006).

Most forms of noise-induced hearing loss are permanent, due to irreparable damage to hair cells in the inner ear (as discussed later in this chapter). Indeed, studies show evidence of irreversible hearing loss manifested in young people who use PLDs. For instance, an audiometric study found diminished hearing in the 3000 to 8000 Hz range in university students who had used PLDs for 1 to 5 years, compared to peers who did not use PLDs (Peng, Tao, & Huang, 2007). Earbud users tend to experience more damaging effects than those using earphones that fit over the ear, as studies show that earbuds allow more ambient noise to leak through, so users' listening levels tend to be significantly higher (e.g., Hodgetts, Rieger, & Szarko, 2007).

Aside from limiting one's use of PLDs and monitoring loudness levels of devices, using earphones that isolate ambient noise can be helpful for volume control (Portnuff, 2016). Knox (2009) also found that adding a simple visual feedback accessory that helps listeners monitor their listening levels on PLDs was effective in reducing listening levels. For general environmental noise or when attending musical events, ear plugs are an inexpensive method for preventing MIHL. More expensive ear plugs can be custom-made for orchestra musicians, for instance to allow low-frequency sounds of a double bass to be heard while blocking higher-frequency sounds of other instruments close by. More public education on MIHL is needed via entertainment channels, health care professionals, college courses, and through corporations that design and manufacture PLDs, headphones, and earbuds.

Duration

Musical tones also differ in duration. Such differences are represented in musical notation by the symbols for whole note, half note, quarter note, and so on (semibreve, minim, crotchet, quaver, etc.), each representing a tone of one half the duration of the preceding note value. However, these are simply the subdivisions of note durations that conveniently organize tones in most Western music. In practice, musical performance involves a lot of finer-grained variations in timing, and our ability to sense these small variations in timing is important to our ability to perceive expressivity in music. (Expressive timing is discussed in greater detail in chapter 11 on performance.)

The duration of the acoustic signal may not solely determine the perceived length of a tone. Schutz and Lipscomb (2007) asked internationally acclaimed percussionist Michael Burritt to play a series of tones on a marimba – using either long graceful gestures or short clipped gestures to strike the keys. Their acoustic analysis revealed that the resulting tones were acoustically indistinguishable in duration. The long- and short-gesture tones were also indistinguishable to listeners hearing an audio recording of those tones. However, participants tended to perceive the sound of the tones to be significantly longer when accompanied by a video showing Burritt playing with long and graceful gestures, and shorter in duration when accompanied by a video showing short, choppy gestures.

Figure 3.1 shows time-lapse images of the videos used in this study, and the videos themselves may be viewed at www.maplelab.net/illusion.[1] Notice that the 'long' stroke traces

Figure 3.1 Time-lapse images of the videos used in Schutz and Lipscomb's (2007) study, which may be viewed at www.maplelab.net/illusion. Captured 200 milliseconds apart, they show that acclaimed percussionist Michael Burritt's striking mallet (held in his right hand) continues moving for a longer time after impact for the 'long' stroke (top row) versus the 'short' stroke (bottom row). Although the difference in post-impact motion does not affect the tone's acoustic duration, it does affect our *perception* of the duration of the tone.

Source: Copyright © Michael Schutz and Scott Lipscomb, printed with permission of Michael Burritt.

a smooth arc after impact, whereas the 'short' choppy stroke stops abruptly. Schutz and Lipscomb concluded that the 'difference' participants 'heard' between tones created by long-slow and short-swift strokes was perceptual rather than physical, caused by 'visual artifacts of the performer's acoustically inconsequential gesture' (2007, p. 896). There is some similarity to the 'McGurk effect,' which can be viewed at https://vimeo.com/200587571,[2] in which watching the lip movements of a speaker alters the listener's perception of the syllable being spoken (McGurk & MacDonald, 1976). In a subsequent study, Schutz and Kubovy (2009) showed that this illusion persisted even when the video of Burritt was reduced to a point-light display consisting of only a single white dot tracking the moving mallet head against a black background, as can be seen at www.maplelab.net/virtualmarimbist. This is just one illustration of how the pure physical signal and the perception of a sound may differ. Other examples, pertaining to pitch perception, will be discussed in chapter 5.

Our perception of tone durations is also circumscribed by the physical limitations of the auditory system. For example, clicks are perceived as if occurring simultaneously if separated by less than 2 milliseconds, as shown in pioneering research by Exner (1875). Similarly, there is a threshold of order: Sound events are heard as separate, but their temporal order is not discernible for separations less than 15 to 20 ms (Hirsch, 1959). Sounds of very short duration may not even be perceived as a tone. For instance, a very short tone burst (2 to 4 cycles in length) is usually heard as a click without a discernible pitch (Pierce, 1992). The higher in pitch the sound is, the longer the tone burst needs to be if there is to be no click. As we shall see in the following section, there is also a temporal (i.e., time-based) dimension to the perception of timbre.

Timbre

Timbre or 'tone color' is the intrinsic and distinctive quality of the sound produced by a musical source such as a certain instrument or a particular human voice. The brilliant sound of the trumpet contrasts in timbre with the mellow, reedy sound of the clarinet. Timbre perception is a very complex topic that poses many methodological challenges to researchers (see Hajda, Kendall, Carterette, & Harshberger, 1997). In fact, the definition of timbre given by the American National Standards Institute is based on timbre being *any* quality of a tone other than its pitch, loudness and duration! For the purposes of this brief discussion, we focus on two aspects: *frequency relationships* (also discussed in chapter 2) and *changes across time*.

Frequency relationships

The quality we refer to as 'timbre' corresponds with the ***Fourier analysis*** of the waveform, which reflects the various frequencies it contains. A pure tone, such as the ringing of a tuning fork, is produced by a sound wave of a single frequency of vibration. However, as mentioned earlier, such a smooth wave would be more characteristic of the sound made by a well-crafted tuning fork, as opposed to the more complex sound of an acoustic instrument. Each tone produced by most musical instruments comprises not only the fundamental frequency, but also its overtone series (as shown in Figure 2.4 in the previous chapter). The particular combination of overtones sounded by an instrument is responsible, in part, for its distinctive tone quality or 'timbre,' and provides important cues for the accurate identification of the sound of musical instruments by both humans and machines (Brown, Houix, & McAdams, 2001; Martin & Kim, 1998).

Figure 3.2 shows the characteristic waveforms of a guitar, violin, vibraphone, and double bass, along with associated power spectra. The waveforms (left column) are best used to see differences in fundamental frequency (pitch); in complex tones like the ones shown in the figure, fundamental frequency relates to the repetition rate of the overall complex vibration

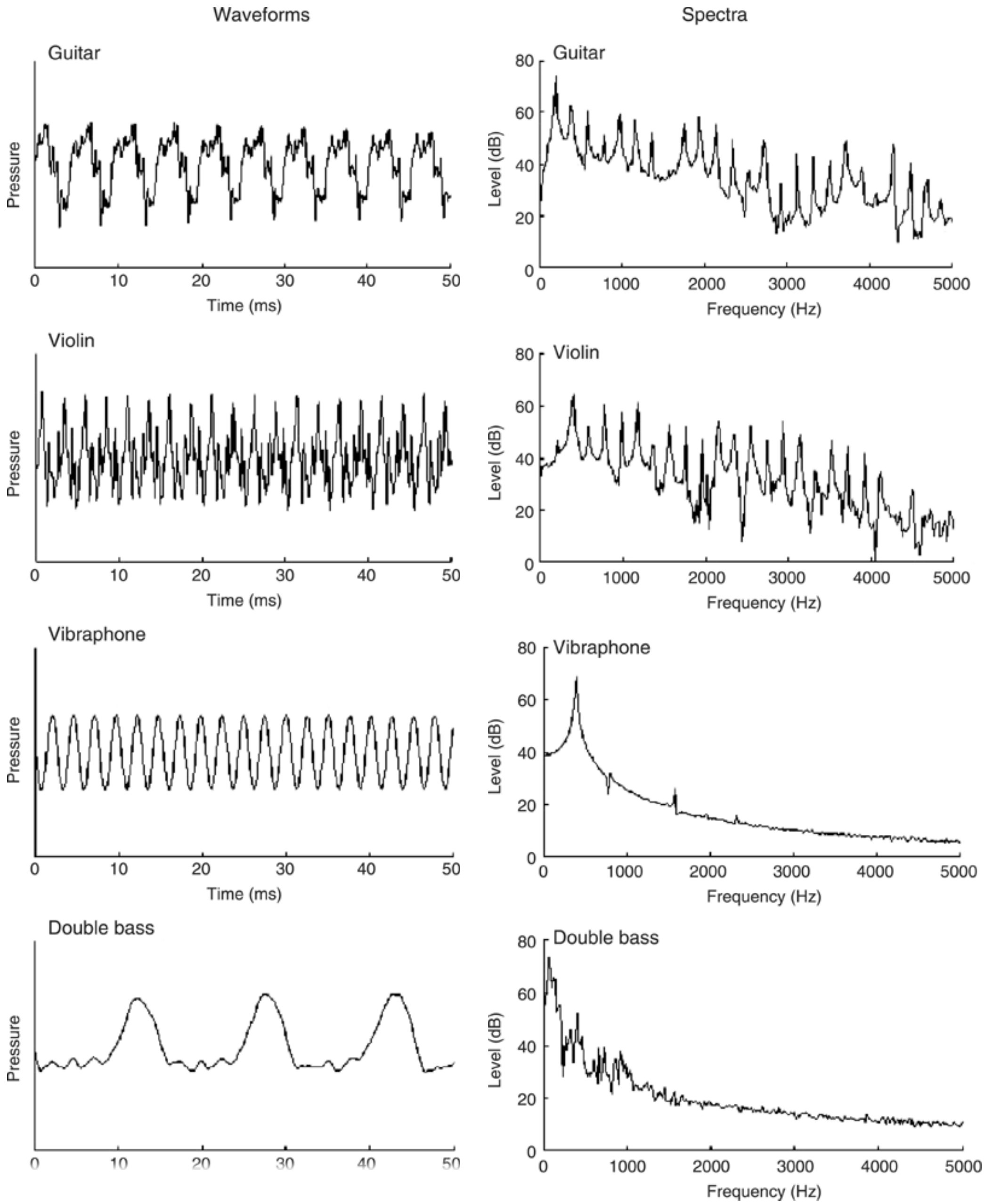


Figure 3.2 Waveforms (left) and power spectra (right) for a variety of musical instruments.

Source: Reproduced from Plack (2014, p. 34) by permission of the author.

pattern. The double bass plays a low pitch, and you can see just three repetitions of its vibration pattern within 50 ms, whereas the vibraphone and violin play higher pitches with associated oscillations that repeat many more times within the span of 50 ms. The ***power spectrum*** (shown on the right) relates mostly to differences in timbre. The power spectra of the two instruments displayed on the top (guitar and violin) have more energy at high frequencies than do the two instruments displayed below. These differences reflect the timbre of instruments: Whereas the violin can produce a more piercing and brighter sound, the vibraphone is more mellow. As discussed in chapter 2, this difference can be quantified using the measure of ***spectral centroid*** (the central tendency across frequencies in the power spectrum), which is lower for the mellower-sounding vibraphone.

Changes across time

Along the temporal (i.e., time) dimension, the ***amplitude envelope*** of a non-percussive musical signal can be divided into three segments: (1) ***attack*** (roughly, from the onset of the sound to the initial achievement of a steady tone); (2) ***steady state*** (the time during which the sound is fairly constant); and (3) ***decay*** (the time it takes for the sound to fade away). For instance, the attack of a bowed violin is quite slow; it takes 1/5 of a second or more before the string is fully engaged in the regular motion that produces the consistent sound of the steady state (Beament, 1997), followed by a gradual decay. By comparison, a cymbal struck by a hard mallet has an almost instant attack, no true steady state, and a very slow decay. This 'percussive' envelope is characteristic of many acoustic instruments. Interestingly, although percussive envelopes are highly common in music, most publications to date in the leading journal of music cognition, *Music Perception*, have used completely flat amplitude envelopes (with no amplitude fluctuations across its duration), which are highly unnatural (Schutz & Vaisberg, 2014).

The 'attack' may play an important role for the identification of musical instruments (for further discussion, see Hajda et al., 1997). For instance, if the initial portion of a recorded tone (such as the first 1/5 of a second of the starting sound of a trumpet) is removed, it is often more difficult for listeners to identify the instrument (Saldanha & Corso, 1964). This may be in part because instruments have distinctive ***attack transients***, or temporary fluctuations in the sound signal that occur during the initial phase of a tone. These 'impurities' in the beginning of the attack help us identify an instrument; for instance, attack transients are largely responsible for the starting 'blat' of a trumpet or the initial gritty sound of a low cello tone.

This is not to say that the 'steady state' and 'decay' are not also important for identification of musical instruments. For instance, Saldanha and Corso also found that instruments with a vibrato tone in the steady state were more easily identified than were those with a non-vibrato tone. This may be because width, rate, duration, and other characteristics of the modulation in the vibrato differ between instruments. For instance, the vibrato in voice and violin is mainly a result of modulation in frequency (pitch), whereas trumpet vibrato mainly stems from modulation in amplitude (loudness) (Geringer, MacLeod, Madsen, & Napoles, 2015). Further, whereas transients seem to provide important cues for identifying the timbre of single musical tones, Kendall (1986) found the steady state to be more important when identifying instruments playing short musical phrases.

Can computers and other machines be programmed to identify musical instruments by sound? In comparison to the large body of work on speaker identification by computers, few published studies have focused on identification of musical instruments. In one study, Judith Brown and colleagues (2001) found that computers performed comparably to humans in identifying four wind instruments. Flute, clarinet, oboe, and saxophone were chosen as these instruments are difficult for many humans to identify. Machines can even be 'taught' to judge 'good,' 'medium,' and 'bad' trumpet tones in a way that corresponds with human brass players' evaluations of the quality of the timbre (Knight, Upham, & Fujinaga, 2011). Machines performed at above chance level on this task, though not with high accuracy.

The anatomy and physiology of the ear

Given the complexity of the neurophysiology of hearing, particularly as it relates to the perception and production of music, it will be helpful to describe the anatomy and physiology of the auditory apparatus (the outer, middle, and inner ear), before describing the processes in the brain through which auditory experience is produced.

The auditory system consists of (1) the external ear, (2) the middle ear, and (3) the inner ear. Its major divisions are shown in Figure 3.3. The ***external ear*** gathers the trains of pressure waves in the air and directs them into the auditory canal. The ***middle ear*** includes the apparatus of tiny bones, or *ossicles*, by which sounds are transmitted to the ***inner ear***, the main organ of which is the cochlea. The *cochlea* is the intricate organ that, with its complex sound detectors, serves to transduce the relevant physical properties of wave trains into neural impulses, which are transmitted through complex networks of neural tracts to various parts of the brain.

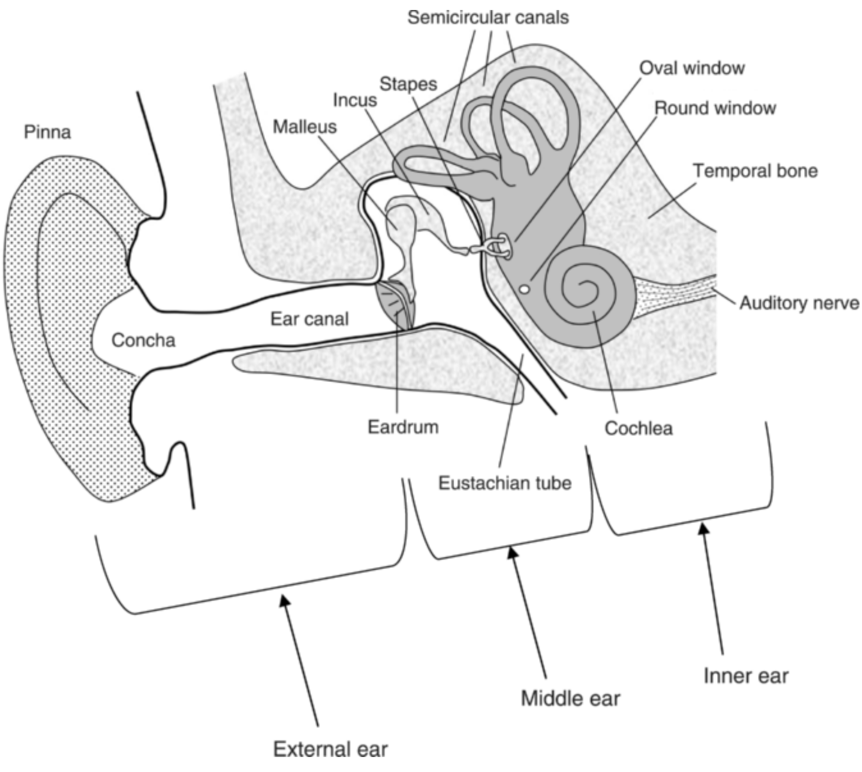


Figure 3.3 The main structures of the ear.

Source: Reproduced from Plack (2014, p. 54) by permission of the author.

The external ear

The external ear comprises the outer ear (or *pinna*) and the *auditory canal*. The ***pinna*** is the external part of the ear (the flap that protrudes from the head) that 'scoops up' the sound waves, directing the incoming vibrations into the auditory canal toward the eardrum. The ***auditory canal***, an S-shaped tube of about 2.5 cm in length, terminates in the *eardrum* or *tympanic membrane* (the membrane that vibrates to the energy and frequencies of the incoming wave trains). The auditory canal acts as a resonator (especially for frequencies ranging in the highest octave of a grand piano), so there is some amplification of the sound waves as they travel down the ear canal.

The pinna plays an important role in ***sound localization***, which refers to the ability to detect where sounds are coming from. The asymmetrical form of the pinna (i.e., a half-moon shape, joined on one side to the head) serves us well in this respect: It allows us to receive more direct sound from sources in front of us than from behind us. If the ear were completely round and affixed to the head by its center, localization of sounds originating in front or behind us would be extremely difficult! Further, the irregularity of the ridges and convolutions of the outer ear offer important cues about the location of a sound, as they create a series of reflected waves before the sounds enter the ear. The pinna also filters frequencies depending on the angle of the sound source. For instance, the irregular form of the pinna selectively transmits more high-frequency components from sources above us than those that are level to the ear, giving us important cues about the elevation of a sound source.

The shape of the pinna varies widely in the population, so sound localization cues that rely on the pinnae are highly idiosyncratic. Indeed, when four participants were fitted with small prostheses that changed the shape of their outer ears and wore them continuously for six weeks, their ability to accurately localize sounds declined! After about six weeks, their performance rose again to baseline levels as they grew accustomed to their 'new ears' (Hofman, van Riswick, & van Opstal, 1998).

Unlike most mammals, human beings cannot move their pinnae toward the direction of sounds; they must move the whole head to achieve this task. Owls, in particular, are extremely accurate at localizing sound sources in part because of their capacity to rotate the head by about 270° (not exactly a complete rotation, as commonly thought). Further, the entire facial ruff (the long feathers on the head and neck) serves a similar function to the human pinna, as these sound-reflecting feathers provide cues as to the source of an auditory event (Knudsen & Knudsen, 1989).

The sound localization cues supplied by the pinnae are referred to as ***monaural cues***, because they only require the use of one ear. Author ST's father suffered sudden hearing loss several years ago and is completely deaf in one ear. However, he can still localize sounds like a person walking up behind him, mainly by relying on cues supplied by the pinna of his single hearing ear. Other forms of localization, such as whether a person is on his right or left, are more difficult. To compensate, he makes more head and body movements and must pay more attention to his surroundings.

In addition to monaural cues, one can also use ***binaural cues***, which involve both ears. Because we possess two ears that are spaced some distance apart, a sound originating from a source closer to one ear arrives at that ear slightly faster –***interaural timing difference (ITD)*** – and/or at a higher intensity – ***interaural level difference (ILD)*** – than at the other ear. For instance, if a sound source is closer to the left ear, the sound has to travel past the head to reach the other ear. The head also acts as an obstacle, absorbing some of the sound and blocking its direct path to the other ear, creating a 'sound shadow.' The resulting sound will then arrive sooner and be more intense at the left ear than the right ear, and the listener perceives the sound as originating from their left.

Sound localization also involves cues that extend far beyond the shapes of the two ears or the distance between them. The entire body, especially the head, shoulders, and chest, plays a role in reflecting and diffracting sound waves, bringing about significant changes in the composition of frequencies that reach the ears. Another factor may have to do with the fact that the phase – that is, the pattern of the crests (high-pressure phase of the sound wave) and troughs (low-pressure phase of the sound wave) of the incoming waves – is different at each ear unless the head is aligned with the track of the sound waves. A full discussion of this effect is covered in volumes focusing on auditory perception (e.g., Handel, 1989, chapter 3).

The eardrum and the transmission system of the middle ear

The middle ear consists of a system of three tiny bones or ***ossicles*** that directly link the tympanic membrane with the *oval window* of the cochlea. These ossicles are illustrated in Figure 3.3, and are referred to as the ***malleus***, ***incus***, and ***stapes***. These titles come from Latin words relating to the shape of each bone: *hammer*, *anvil*, and *stirrup*, respectively. The malleus is attached to the inner surface of the tympanic membrane, and is connected to the incus by a flexible joint. The incus is then connected to the stapes. The stapes is attached to the flexible oval window of the cochlea. Thus, the movement of the tympanic membrane sets these three bones in motion and leads to a pattern of vibrations on the oval window that represents the sound wave.

The middle ear is designed to preserve information in the sound wave as it is transferred from one medium (air) to another (fluid, for the inner ear). The problem the middle ear is designed to solve is that fluid is of a higher density than air, and so a similar amount of energy will be less effective in moving fluid than in moving air (think about waving your arms in the air versus through water). Thus the middle ear accomplishes ***impedance matching***: that is, the preservation of energy across two media that differ in their levels of resistance. The ***impedance*** of any medium is defined as the ratio of the pressure imposed on a medium (stapes pushing on pushes oval window) to the resulting flow within the medium (fluid movement in the cochlea). The air – the medium of transmission for sound in the external environment – is of low impedance; that is, it has little intrinsic resistance to motion. However, the cochlea is filled with fluid, which is of high impedance. The angles formed by the ossicles form a lever-like mechanism that serves to amplify vibrations in the sound wave. This process is further served by the difference in sizes between the tympanic membrane (larger) and the oval window (smaller).

The middle ear is also designed to keep loud sounds out of the cochlea. Two muscles (*tympanic* and *stapedius*) attached to the ossicles contract in response to very loud sounds, attenuating sounds below about 1000 Hz. This ***acoustic reflex*** protects us from loud incoming sounds, but also occurs when we talk or sing (even at moderate levels of 65 or 75 dB), shielding us from the effects of exposure to our own voices. The reflex, however, is too slow to protect the inner ear from sudden bursts of loud sounds (which is why ear plugs are especially essential for percussionists).

The inner ear: The structure of the cochlea

The main structure of the inner ear is the cochlea. The ***cochlea*** is a coiled, snail shell-like arrangement of three parallel tubes – the *vestibular canal* (*scala* is used for canal in

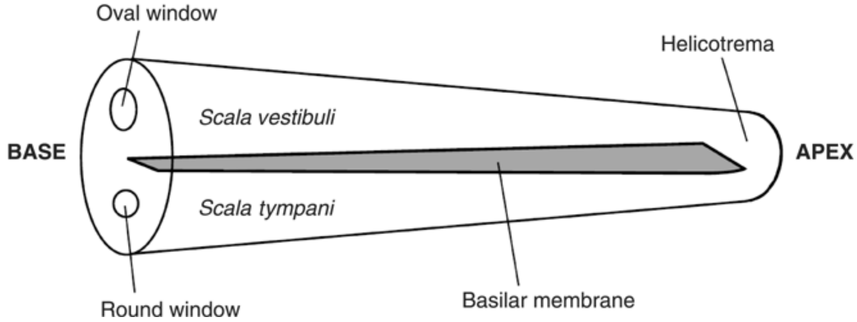


Figure 3.4 The structure of the unrolled cochlea. This figure shows the basilar membrane, but not other structures of the median canal.

Source: Reproduced from Plack (2014, p. 56) by permission of the author.

Figure 3.4), the *median canal*, and the *tympanic canal* – filled with an incompressible fluid (*endolymph* in the median canal, *perilymph* in the others). If the cochlea were unrolled and viewed from the side (as in Figure 3.4), it would measure about 3.5 cm in humans. The median canal is separated from the others by two flexible membranes, the *basilar membrane* (which separates median from tympanic) and *Reissner's membrane* (which separates median from vestibular). The median canal contains structures that allow auditory perception to occur, and we focus on these structures next.

On the surface of the ***basilar membrane***, within the median canal, the entire neural apparatus for sound detection is laid out in the ***organ of Corti***, shown in Figure 3.5. The perspective of Figure 3.5 is a 90° rotation of the perspective used in Figure 3.4, as if we were peering down the length of the cochlea, and the level of detail is much higher, as we are focusing on a specific structure in the median canal – which in Figure 3.4 was reduced to the basilar membrane, reflecting the small size of that canal. The organ of Corti contains bands of sensitive ***hair cells***, in two rows, laid out in an inner and outer array (shown on the left and right of Figure 3.5). An astonishing video of the responses of a single outer hair cell recorded by Dr. Jonathan Ashmore can be viewed here, along with a brief description by Schnupp, Nelken, & King (2011): https://auditoryneuroscience.com/ear/dancing_hair_cell.

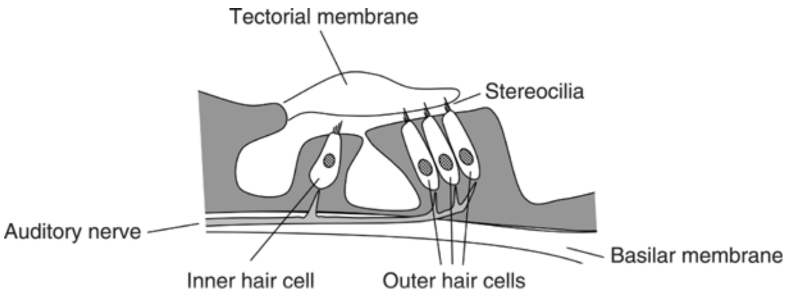


Figure 3.5 A cross-section of the cochlea rotated 90°, showing the organ of Corti, including inner and outer rows of hair cells.

Source: Reproduced from Plack (2014, p. 57) by permission of the author.

Although both rows of hair cells are affected by the incoming sound wave, the outer hair cells are particularly susceptible to damage caused by loud sounds (as described in our earlier discussion of MIHL). When outer hair cells are damaged, soft sounds may not be perceptible, while other sounds may seem abnormally loud (a condition known as 'loudness recruitment'). The ability to perceive high-frequency sounds is also diminished, affecting perception of timbre and tonal nuances of music. As hair cells do not regenerate, the effects are irreversible and cumulative.

How does sound travel through the inner ear? Pressure is transmitted by the stapes, fixed to the flexible membrane in the ***oval window*** of the vestibular canal. This results in pressure waves that are induced in the fluid in the cochlea by the action of the stapes on the membrane of the oval window. Since the vestibular fluid is incompressible, the pressure of the wave must be relieved. This is achieved by the flexing of the membrane in the ***round window*** (the second aperture in the tympanic canal) when the pressure wave reaches it. (The main structures described here are labeled in Figure 3.4 and in our main illustration of the ear in Figure 3.3.) The main mechanism of pitch detection depends on the induction of a traveling wave in the basilar membrane. In order to understand this process, we need to know more about how the cochlea both *transduces* and *encodes* frequency information in the sound wave. Thus, transduction and encoding are discussed in the next two sections.

Neurophysiology of the cochlea as transducer

A complex process occurs during sensation: Energy of one form (like a sound wave) must be converted into an electro-chemical energy (a neural impulse). This process of converting one form of energy to another is called ***transduction***. This transformation must preserve certain *relevant* features of the original process. For instance, when pressing keys while typing at a computer keyboard, we are involved in the process of *transducing* mechanical impulses (movements of keys) to electrical signals in the device. The device has to retain information about what key was pressed even though the physical signal has changed significantly.

When we hear music, the pattern of physical disturbances from the original vibration (e.g., vibrating vocal cords, in the case of singing) is generally retained through the ear. But following the wave's arrival in the cochlea, transduction must occur in order for the information to reach the brain, since the brain only 'understands' electro-chemical signals. For music, the relevant physical information can be reduced to frequency (heard as pitch) and amplitude (heard as loudness). The form of this information is that of induced pressure waves in the perilymph (fluid) surrounding the median canal. Somehow, all the energy resulting from colliding fluid molecules in the inner ear must be converted to neural signals that are reconstructed in the brain so that we can experience the intended musical message being performed.

As the basilar membrane is progressively displaced by the movement of fluid in the cochlea, hair cells embedded in the *organ of Corti* (Figure 3.5) push against the dense and positively charged endolymph that fills the median canal. When this happens, hair cells bend and a small opening at the base of the hair cell is opened, and positively charged potassium ions flow inward. This alters the potential difference between the inner and outer regions of the hair cell, thus creating an action potential. At this point, sound energy is transduced into a nerve impulse, though only the inner hair cells convey information about sound vibrations. A complex tone with three frequencies of 200 Hz, 400 Hz, and 800 Hz, for example, will lead to a traveling wave in the basilar membrane with three regions of relatively maximal oscillations. As a result, sets of hair cells at three corresponding places on the membrane will respond, thus mapping sound frequency onto the cochlea tonotopically (as will be explained later in the chapter).

Encoding of pitch in the cochlea

Transduction involves a dramatic change. The sound wave is converted from a continuous pattern of vibration to a series of discrete action potentials distributed across many neurons. It is important that this change does not distort the incoming information too greatly. A sensory system must preserve information from the outside world even as that information is transduced, a process called ***encoding***. The cochlea is designed to encode information about frequency in sound, which is central to the perception of music.

Our understanding of the neurophysiology of the cochlea is based on the results of the extensive research program undertaken by Hungarian physiologist Georg von Békésy (1899–1972), which earned him the Nobel Prize for Physiology or Medicine. From von Békésy, we learned that areas of the cochlea respond differentially to sounds depending on their frequencies. Since then, two dominant theories have emerged to explain how the frequency of a sound wave is encoded as pitch, one for sounds of frequency more than 5000 Hz and one for sounds of lower frequencies. These are often referred to as the place theory and time theory.

According to the ***place theory***, the relevant parameter is *where* on the basilar membrane the maximum displacement occurs. This location is populated by a certain row of inner hair cells that are maximally stimulated by the displacement of the membrane, perhaps amplified by changes in the shape of the corresponding outer rows of hair cells. The place theory suggests that the basilar membrane responds in a way that separates each sinusoidal component of a complex tone, performing a sort of Fourier analysis. In such a framework, pitch perception may be determined by the 'place' associated with the fundamental frequency. However, place mechanisms cannot account for the aforementioned phenomenon of *residue pitch*, for which there is no 'place' corresponding to the fundamental that is heard as the pitch. This problem has led some to propose that pitch perception results from temporal properties in the overall spectrum of a complex tone, as in the theory we describe next.

According to the ***time theory***, the relevant parameter is the *time* between successive phases of the displacement of the basilar membrane by the energy of the traveling wave. A sine wave makes a complete cycle from 'crest' to 'trough' and back again to a 'crest.' The higher the frequency, the shorter the time between each half phase of the cycle, and so the more rapidly the bursts of firings in the auditory nerve will follow one another. With respect to complex tones, the overall pattern of firing can be used to identify the fundamental even if the fundamental frequency is absent and one must rely on residue pitch instead (cf. Cariani & Delgutte, 1996). However, the time theory also runs into problems. For vibrations of less than 5000 Hz, the auditory nerve bundle fires for only half the cycle of stimulation, but for vibrations above 5000 Hz, it fires continuously. If the auditory nerve fires continuously, this source of pitch information disappears since there is no time difference between successive levels of neural activity – for example, successive 'crests' of the energy wave.

In sum: Both 'place' and 'time' theories ultimately seem necessary to account for the perception of pitch (Cariani & Delgutte, 1996), because both mechanisms contribute to pitch perception and have different limitations. Place information is the major factor in encoding pitches above 5000 Hz (Pickles, 2008, p. 273), whereas time is the major factor for pitches below 50 Hz (Warren, 1999). The range of pitches that remain – which includes most of music – probably draws on both codes (Plack, 2014).

So far, we have described in outline how the outer, middle, and inner ear play their parts in an auditory transducer. Pressure waves in the air are represented by impulses from the neural apparatus of the pair of cochleae. These impulses pass into the brain, and are eventually experienced by listeners as sounds, and some of these sounds as music. The next step will be to describe the first stages of auditory processing in the brain. More complex processes related to music perception are described in subsequent chapters.

Auditory pathways to the brain

The neural pathways from cochlea to auditory cortex

The pathways leading from ear to cortex are complex and tangled. Though a great amount of physiological research has been done on these connections, the significance of these findings for the experience of complex sound patterns such as music is not yet clear. A remarkable theme in these pathways, in contrast to visual pathways, is the degree to which auditory pathways provide both ***ipsilateral*** (same side) and ***contralateral*** (opposite side) connections from ear to brain. In other words, there is much more sharing between the left and right in our auditory system than in our visual system. Some neurons receive excitatory input from one side and inhibitory input from the other side (Plack, 2014). This kind of setup may relate to the fact that spatial relationships in hearing often require a comparison between inputs to the left and right ears, such as when listening to music through headphones.

The auditory cortex

The ***auditory cortex*** is a bilateral structure in both left and right temporal lobes, consisting of three layers of neurons: the *primary*, *secondary*, and *tertiary* regions. The layers are arranged in concentric circles, the ***primary auditory cortex*** occupying the center, surrounded by the cells of the secondary and tertiary regions.

The primary auditory cortex consists of columns of neurons that appear to be ***tonotopically*** arranged; that is to say each cell in a column is tuned to the neuronal signal from a particular frequency in the auditory input into the cochlea. Tonotopic mapping matches a location in the primary auditory cortex to a location on the basilar membrane of the inner ear. Thus the organization of the primary auditory cortex matches that of the tonotopic layout of the bands of hair cells along the basilar membrane in the cochlea. Higher frequencies are recognized towards the central area of the primary auditory cortex, and lower frequencies nearer the boundaries.

The significance of tonotopic organization is that the auditory system represents pitch in a manner analogous to a piano, in that there is a spatial representation of pitch. Furthermore, the tonotopic organization is laid out such that the brain encodes pitch in absolute terms (e.g., there is a specific region associated with C4 or the fourth highest C on a grand piano). However, if primary auditory cortex is organized in absolute terms, the ability to perceive pitch in this way, called ‘absolute’ or ‘perfect’ pitch ability, should be more widespread than it is. We discuss the intriguing phenomenon of absolute pitch perception further in chapter 5.

Adjacent to each primary auditory cortex are the ***secondary*** and ***tertiary auditory regions*** in concentric layers, referred to as the ‘*belt*’ and ‘*parabelt*.’ These are illustrated in Figure 3.6. In this figure, the fold containing the auditory cortex (called ***Heschl’s gyrus***) has been cut and opened. This figure shows the brain of a macaque monkey, which is structurally similar to that of a human brain with respect to the auditory cortices. As the names ‘secondary’ and

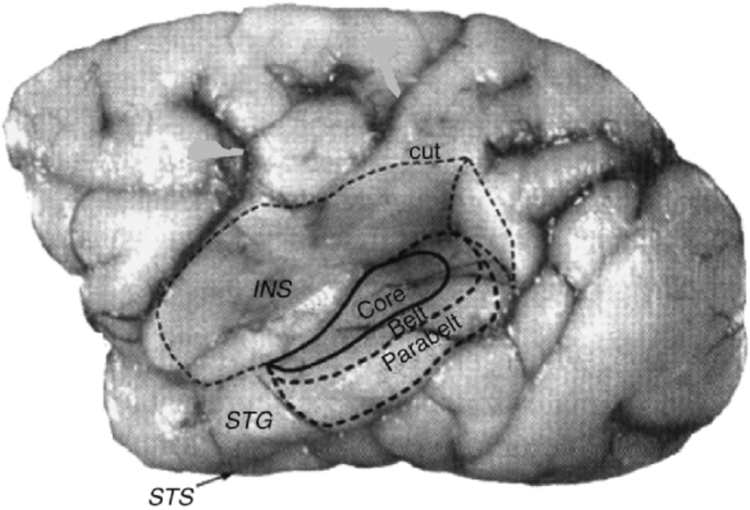


Figure 3.6 Secondary and tertiary auditory regions in the macaque monkey.

Notes: *INS* = insula, *STG* = superior temporal gyrus, *STS* = superior temporal sulcus.

Source: Kaas & Hackett (2005, p. 11), reprinted with permission from Lawrence Erlbaum.

'tertiary' imply, these regions appear to be important for more complex aspects of auditory experience. Experiments have shown that the cells in these regions do not respond at all, or if they do so, only weakly, to pure tones. Instead, these areas respond to complex tones in ways that distinguish 'what' from 'where' similarly to higher processing areas of the visual system (Rauschecker & Tian, 2000). This architecture may have practical implications for music listening in that, for instance, we can identify the sound of an instrument independently of its location in space (e.g., during a Broadway musical, in which singers frequently move around in relation to each other while singing).

Cochlear implants and perception of music

Thus far, we have focused on hearing in the typical population. In this final section, we turn to the topic of how individuals with cochlear implants (CI) may experience music. Cochlear implants first became a predominant method of addressing profound deafness in many parts of the world in the 1980s. Since then, several hundred thousand devices have been implanted in individuals, ranging in age from infants under 12 months to recipients over 100 years old! ***Cochlear implants*** are devices that are surgically implanted within the inner ear, coupled with an external sound processor that is placed behind the outer ear (pinna) that stimulates the auditory nerve, using an electrical current to transmit signals to the brain. Unlike a hearing aid, which amplifies sound, CI gives the user a representation of sound, so it may take a formerly hearing individual some time to adjust to a new way of experiencing sound.

Before reading on, please listen to this striking simulation of what Beethoven's *Kreutzer Sonata* might sound like if listening through a CI device: https://auditoryneuroscience.com/prosthetics/music, found at the companion site to *Auditory Neuroscience* (Schnupp, Nelken, & King, 2011). As with all simulations, the demonstration provides only an approximation of what CI users might hear, and that individuals who use CI vary widely in their experience of sound. However, the main point of the simulation is clear: Although helpful for facilitating comprehensibility of speech, current models for cochlear implants are not conducive for the perception of music, which comprises much more complex auditory information.

In tonal tests, adult CI users tend to have great difficulty identifying melodies (especially melody line alone) and singing back pitch patterns (e.g., Nakata, Trehub, Mitani, & Kanda, 2006). The time-based aspects of timbre, such as the amplitude envelope, are mainly preserved. However, much spectral information is lost (such as the characteristic overtones of sounds produced by instruments and the relative strength of those overtones). Thus, CI users often have difficulty recognizing instruments, not only within but across families – for instance, mistaking a string instrument for a wind instrument (Gfeller, Witt, Woodworth, Mehr, & Knutsen, 2002).

Compared to tonal or melodic tests, performance of CI users tends to be stronger on rhythmic tests (Innes-Brown, Marozeau, Storey, & Blamey, 2013) and is sometimes almost as good or as good as in individuals with normal hearing (Drennan & Rubinstein, 2008). For example, Phillips-Silver and colleagues played a Latin *merengue* dance piece entitled 'Suavamente' to adult CI users. When the original *merengue* dance music (with prominent vocals and various instruments) was played, or when a version was played on a piano, they were unable to synchronize their movements to the music. However, the adult CI users could easily find the beat and move to the music when the strongly rhythmic *merengue* music was played mainly on unpitched drums (Phillips-Silver et al., 2015).

Children with CI and musical engagement

CI users who lost their hearing at a later age and were adults at the time the device was implanted often express a strong dislike of music (Gfeller et al., 2000), particularly if they have a strong memory of their experience with music prior to the CI. They often say that they no longer recognize the sounds as 'music.' On the other hand, children whose experience of music is largely limited to perception via CI tend to report more enjoyment and engagement with music (e.g., Mitani et al., 2007; Vongpaisal, Trehub, & Schellenberg, 2006). Unlike adults who lost their hearing after many years of experience with music, CI users who are congenitally or prelingually deaf may approach music differently from the beginning. Without the disappointing sense that their perception of music has been degraded, children can focus on what their experience of music offers. For instance, their focus on timing cues inherent in music, as opposed to pitch, enables children to engage with music by moving in time to it, and to participate in motor synchrony – such as clapping or moving in time with other people (Vongpaisal et al., 2006).

Using their greater sensitivity for temporal (i.e., time-based) patterns to compensate for their lack of acuity of pitch patterns allows CI users to appreciate various facets of musical experience. For instance, Hopyan and colleagues found that children and adolescents (aged 10 to 15 years) who use CI could differentiate between excerpts of classical piano music that expressed happiness or sadness, based mainly on the tempo (fast, slow) of the music (Hopyan, Manno, Papsin, & Gordon, 2016), whereas the comparison group of peers with intact hearing relied more on mode (major or minor). As discussed further in chapter 14, tempo and mode serve as important cues for perceiving emotion in music for both adults and older children, with fast tempo and major mode usually conveying happiness, and slow tempo and minor mode expressing sadness. CI users may also draw on their greater sensitivity to temporal patterns than melodic patterns to facilitate song recognition.

Improving the musical experience for CI users

Attending programs designed for CI recipients or even mainstream methods of music training such as Yamaha music classes may improve pitch perception, recognition of melody, and recognition of timbre. In one study, adult CI users participated in 6 months of a musical ear-training program consisting of weekly private lessons focusing on singing, playing, and listening. After 6 months, adult CI users who received ear-training performed significantly better on tests of timbre (including identifying instruments by sound), melodic contour, and rhythm compared to a control group of adult CI users who did not receive music training (Peterson, Mortensen, Hansen, & Vuust, 2012). Many researchers especially emphasize the importance of starting music training at an early age for children with CI (Phillips-Silver et al., 2015).

The growing body of research on the musical capacities of CI users is giving us insight into CI users' experience of music, and advances in cochlear implant design that will accommodate music remain an important goal. In the mean time, some researchers are working on re-engineering existing music in ways that are tailored to CI users' preferences based on knowledge gleaned from research findings, such as reducing music to fewer elements, providing a strong beat and rhythm, and shortening the length of the tones (less reverberation) to make them more distinct and enhance horizontal clarity. For instance, Kohlberg and colleagues found that reducing the complexity of a song by stripping down the instrumentation to fewer musical elements enhanced CI users' enjoyment of the piece (Kohlberg, Mancuso, Chari, & Lalwani, 2015).

Coda

In chapters 2 and 3, we discussed various properties of sound and how sound is perceived in a musical context. The function of the auditory system is to serve as the neural basis of the processes that bring about a conversion from the physical properties of material vibrations, to patterns of electrical and chemical events in the brain, to the sounds we hear as music. Remarkably, this process allows us to identify the sound of a violin or a cello, to locate where each sound is coming from within an ensemble, to detect subtle changes in loudness and timing in an expressive performance, and to be sensitive to the nuances of each tone. Having summarized the basic neurophysiology of hearing, we now turn to the broader neural bases of music perception and performance in the next chapter.

Notes

1 Links to the MapleLab site included with permission of Michael Schutz.
2 The video of the McGurk effect demonstration was produced by Mark Shevy especially for this book.

4 Cognitive neuroscience and the music–language link

It is clearly a time of excitement about the prospects of cognitive neuroscience, in particular neuroimaging. The popular press reports about the neuroimaging studies of music listening have yielded some particularly strong claims, such as 'Listening to music lights up the whole brain' (Suomen Akatemia, 2011). Although this statement is largely misguided – activity can always be found across the entire brain – musical behaviors do stand out with respect to the way in which areas across the brain are involved in music. More important, such broad claims can obscure the fascinating nuances that lead to such widespread activity in the brain. For instance, the scientific report leading to the quote above revealed that different features of music activate different systems in the brain, leading to widespread yet selective patterns of brain activity (Alluri, Toivianien, Jääskeläinen, Glerean, Sams, & Brattico, 2012). Thus, although music may not be distinctive in its use of the 'whole brain,' the intricate coordination of areas across the brain in real time differentiates music from many other behaviors.

This chapter is designed to serve as an introduction to these subtle nuances in brain function, as well as a discussion of the techniques that lead to claims about the brain 'lighting up' (which in fact does not really happen). A full understanding of the neurological activity that underlies the experience and performance of music comes from the study of ***cognitive neuroscience***, the relationship between the brain and thought, involving studies of brain anatomy as well as neural function, or what goes on in the brain. Neuroscience is a broad field that includes many subfields, all of which focus on the brain. It includes traditional ***neurology*** (the study of patients who suffer from brain damage) ***neurophysiology*** (the study of the nervous system on a fine-grained, often cellular level), and ***brain imaging*** (using specialized techniques like functional magnetic resonance imaging to visualize online brain activity). In order to keep the present chapter centered on the psychology of music, we have also interwoven the discussion of a critical research area that has been strongly influenced by neuroscience: the association between music and language. We begin by discussing why the association between music and language is a critical area of research.

The relationship between music and language

Imagine hearing a friend hum the first 'Da-da-da DUM' motif from Beethoven's fifth symphony, and the experience of suddenly recognizing this familiar theme. Now imagine hearing the same friend telling you, 'The Beethoven CD is in the cabinet to your left.' In both cases, your brain has to process auditory information from your friend's voice, compare this sound pattern to information stored in your memory, and make sense of the meanings associated with the sound pattern. However, in many respects these two examples are very different. Whereas we call the first sound pattern music, we call the second one speech or language.[1] These two forms of expression, or 'domains' as they are often called, typically differ in their timing, with music proceeding more slowly than speech, the stability of pitch within notes versus syllables with music being more stable than speech, and whether individual units (e.g., words or notes) refer to specific things in the world (Patel, 2014; Zatorre & Baum, 2012).

Furthermore, whereas the importance of language to our species' survival is clear, the importance of music for survival is highly debated (Cross, 2003, 2005; Huron, 2003; Pinker, 1997). That is, people are interested in whether music is an ***evolutionary adaptation***: an inherited trait that exists because it benefited the survival of one's ancestors. If music and language recruit common brain structures, it is likely that both behaviors originate from a common behavioral precursor in other species (cf. Brown, 2000). If they recruit distinct structures, it makes sense to consider their evolutionary roles separately.

Thus a major point of debate in research on music and language concerns the degree to which these behaviors rely on the same areas of the brain. Evidence to date is mixed, which suggests that some kind of extreme perspective, either complete independence or complete integration, is unlikely to hold true (cf. Patel, 2008). As such, most studies explore the degree to which language and music are integrated. In practice, it can be difficult to determine with certainty whether there are fixed links between brain structures and behavioral abilities. Often it is thought that structure X governs ability Y only because it is not yet known that structure X also governs ability Z. As we will see, neuroscientists' views on which brain areas are specialized for language have changed over time. Moreover, people have come to see the relationship between brain and behavior as more flexible than had been previously thought (e.g., McMullen & Saffran, 2004).

The relationship between music and language ties into a fundamental question concerning how the brain works. Specifically, to what extent are certain behaviors associated with particular areas of the brain? It is common in neuroscience to read that a certain cognitive characteristic (e.g., pitch perception) is governed by neural tissue at a certain location (e.g., primary auditory cortex). This characteristic is known as ***localization of function***. The degree to which functions are localized has important implications for whether the brain is ***modular*** (Fodor, 1983), consisting of independent components that accomplish tasks in an efficient, compartmentalized manner. An alternative view is that the brain functions like an integrated network of connections. This view is often referred to as the ***connectionist*** perspective (McLeod, Plunkett, & Rolls, 1998).

Some anatomical features

The brain is a dauntingly complex organ, comprising billions of cells that can form approximately 100 trillion connections. In order to make sense of our brains, it is common to distinguish brain ***regions*** based on anatomical features. These distinctions exist at varying levels of specificity. A neurophysiologist may focus on ion channels in a cell, or specific folds (gyri) and bumps (sulci) at the brain's surface. Our introduction focuses on a broader (less detailed) scale. For the purpose of illustration, in this section we consider brain areas that are active in a pianist who is sight-reading notation while accompanying a singer.

Figure 4.1 shows lobes of the left hemisphere, which is a common way to segment brain areas on a very broad scale. The divisions don't come from the anatomical features of the brain, but are instead based on the bones that surround the brain (i.e., the frontal, parietal, temporal, and occipital lobes are named for the corresponding cranial bones that underlie them). Each lobe takes on a set of specialized functions, with some important divisions across different lobes.

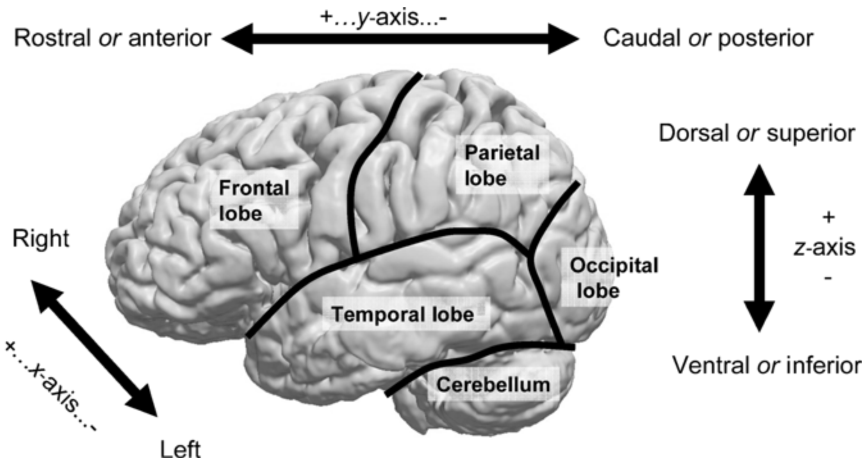


Figure 4.1 Lobes of the neocortex and the cerebellum, along with coordinate labels. The left hemisphere of author PQP's brain is shown.

Source: Brain image scans produced in collaboration with the Buffalo Neuroimaging Analysis Center (Robert Zivadinov, MD, PhD, Jennifer L. Cox, PhD), www.bnac.net. Used with permission from BNAC.

For instance, the ***temporal lobe***, which is particularly critical here, plays a primary role for auditory perception (though it has other roles as well). By extension, the temporal lobe is necessary for the perception of both music and language. The temporal lobe of the pianist in our example will play a vital role in her ability to listen to and synchronize with the singer, as well as process auditory feedback from her piano performance.

Another important lobe in the engagement of performing a piece of music is the ***frontal lobe***, which (among other things) is involved in fine motor control. When our accompanist moves her fingers to play the piano, there is a part of her frontal lobe that directly controls these fine-grained movements. Likewise, there are sections of the frontal lobe designated for control of the tongue, lips, and larynx that the solo singer uses in this performance. As we can see, music performance typically involves the intricate coordination of various brain areas involved in perception and motor control (see chapter 11 for further discussion).

A second way to segment the brain broadly involves dividing it into left and right ***hemispheres***. In general, each hemisphere is *homologous*, meaning it contains structures that match the other hemisphere. At the same time, there is evidence that the right and left hemispheres may perform distinct, specialized tasks. Such differentiation is referred to as ***lateralization of function***. One example of lateralization of function, which we will discuss in more detail later, is the fact that the left hemisphere generally plays a dominant role for language processing, whereas the right hemisphere dominates for processing music. Thus, when the pianist in our example listens to the solo singer, the melodic information from the singer may largely be processed in her right hemisphere, while meanings of the lyrics may be processed in her left hemisphere.

Figure 4.1 also illustrates terms used to identify the location of brain areas in space. The terms used for identifying relative locations in the brain were originally designed with a four-legged creature in mind. This becomes problematic because a quadruped standing on all fours positions its head so that the jaw is roughly parallel to the trunk, whereas it is closer to perpendicular for us. Thus, the term for brain regions toward the bottom are called ***ventral***, referring to the 'belly-side,' and brain regions toward the top are called ***dorsal,*** referring to the 'back.' Dorsal areas are sometimes referred to as ***superior*** to (meaning 'above,' not 'better than') those that are below them. Areas toward the front of the brain may be called ***anterior*** or ***rostral*** (or sometimes simply 'frontal'), and areas toward the back of the brain may be called ***posterior*** or ***caudal***. Again, consistent with the stance of a four-legged creature, rostral refers to the snout or beak, and caudal to the tail. The ***lateral*** (left/right) positioning of brain areas, fortunately, takes the common English words 'left' or 'right,' and areas toward the center are referred to as ***medial***.

Dividing the brain into lobes and hemispheres is useful because these anatomical divisions are associated with different functions of the brain. Even so, we need to subdivide the brain into smaller areas to better understand the relationship between music and language. Figure 4.2 indicates some critical brain regions that will be featured throughout this book. Note that these regions are still based on anatomical divisions (cell composition and folding patterns in the brain, as well as division into lobes). However, in order to clarify the significance of these areas, we discuss functional associations with these regions. It is worth noting that the highlighted areas in Figure 4.2 show how it is a mistake to associate brain functions simply with lobes or hemispheres. For instance, although it is true (as we discussed earlier) that the temporal lobe is important for auditory perception, Figure 4.2 shows that this fact only holds for the superior and medial regions of the temporal lobe. The inferior regions of this lobe are used for visual categorization, including face perception. Thus, the pianist in our example would use the temporal lobe both to recognize the face of the singer she accompanies, in addition to processing the pitches of the singer's voice.

Most areas highlighted in Figure 4.2 are associated with the lobes identified in Figure 4.1, but there is one important area in Figure 4.2 that is not in the cortex: the ***cerebellum***. The cerebellum, which is an area of the hindbrain, is sometimes overlooked in neuroanatomical discussions (early brain scans often left it out), but is included in our discussion because

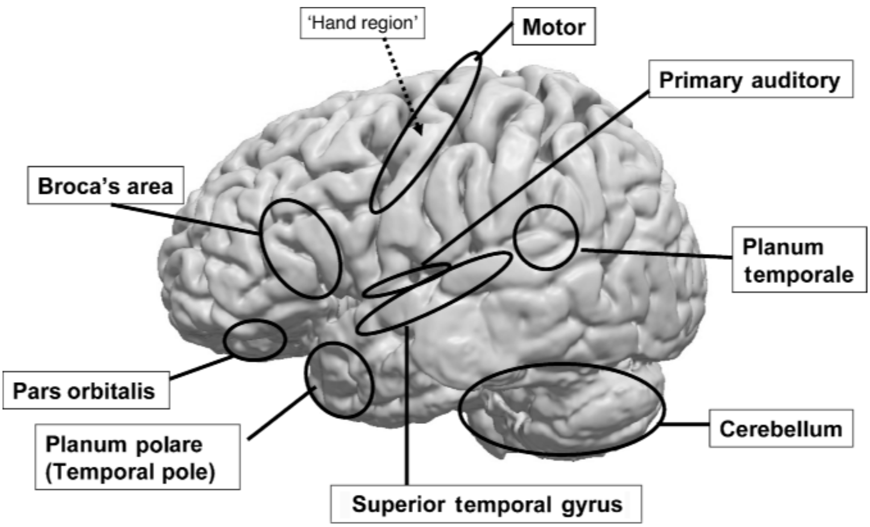


Figure 4.2 Surface brain regions of the associated with music. The left hemisphere of author PQP's brain is shown. Arrows are used to indicate regions on a more localized scale.

Source: Brain image scans produced in collaboration with the Buffalo Neuroimaging Analysis Center (Robert Zivadinov, MD, PhD, Jennifer L. Cox, PhD), www.bnac.net. Used by permission from BNAC.

of its significance to many aspects of musical experience and performance, like the ability of the accompanist in our example to produce regular rhythmic movements (chapter 6). Of course, we should not take the divisions in Figure 4.2 too literally. Every brain's anatomy is slightly different, and we are still learning how the brain works. Regions highlighted in Figure 4.2 should thus not be considered as absolutely precise, but rather as approximations to the brain areas that figure in musical behaviors.

Along with the problem of individual differences, representations like Figure 4.2 can obscure the intricate coordination that exists within an individual's brain. A particular function that a brain area serves – for example, perceiving a pitch – may be accomplished by a specific area, by several areas working together, or within a small subsection of one area.

For instance, although it is true that the temporal lobe in general plays a critical role in auditory perception, more specific divisions within this lobe are important to consider. For instance, when pitches from the singer initially reach the ears of our accompanist, the first brain area to make sense of the pitch content is the ***primary auditory cortex*** (see Figure 4.2 and chapter 3), which is a smaller area within the temporal lobe. The primary auditory cortex is buried in one of the brain's many folds in the temporal lobe, an area called ***Heschl's gyrus***. Pitch processing in this area is thought to involve basic initial processing – hearing the sound as having some kind of a pitched quality, though not necessarily its relationship to a melody. Interestingly, the primary auditory cortex of a highly trained person like our accompanist is probably different in cell composition from that of a musically untrained listener in the audience. Her auditory cortex probably has a higher volume of gray matter than the audience member (Schneider, Sluming, Roberts, Bleeck, & Rupp, 2005).

Following this first stop, the auditory signal reaches many other areas, the first of which were described in chapter 3. Importantly, the stops that the signal takes after the primary auditory cortex reflect more complex aspects of listening. In other words, the auditory cortex helps our accompanist process the basic dimensions of sound, such as pitch content, whereas other areas (such as the superior temporal sulcus, a fold inferior to the auditory cortex) help her categorize the role of these sounds in the melody. A critical question for the music–language debate is whether pitch processing in the auditory cortex proceeds differently for spoken pitch (e.g., rising at the end of a sentence to convey a question) versus musical pitch. We consider this question more in the next section.

Although a good deal of our focus will be on the neocortex and areas in Figures 4.1–4.2, areas in 'lower' regions of the brain are also critical to many musical functions. A few key areas are shown in Figure 4.3. For instance, another area that is important for timing, particularly following the beat, is the ***basal ganglia*** (see chapter 6 for further discussion). Both the singer and pianist in our example rely on this structure to synchronize. This same structure, along with the ***hypothalamus***, is also involved in the strong positive emotional responses that the music elicits in audience members (and also the performers), while moments at which the music can sound surprising and dissonant may elicit responses in the ***amygdala***, which is involved in fear responses. Finally, memory for music (which plays some role, even in sight-reading) relies strongly on the ***hippocampus***.

As mentioned before, the structure of the brain is not determined simply by clusters of cells in an area, but also by connections between these cells. These distinctions are critical to questions concerning whether modularity or connectionism explains how the brain works, as mentioned earlier. Recently, imaging technologies have allowed us to understand the brain not just as a set of anatomically divided regions, but also as a complex system of connections. A certain application of magnetic resonance imaging measures the diffusion of water molecules in the brain. The technique is called ***diffusion tensor imaging (DTI)***. Its logic

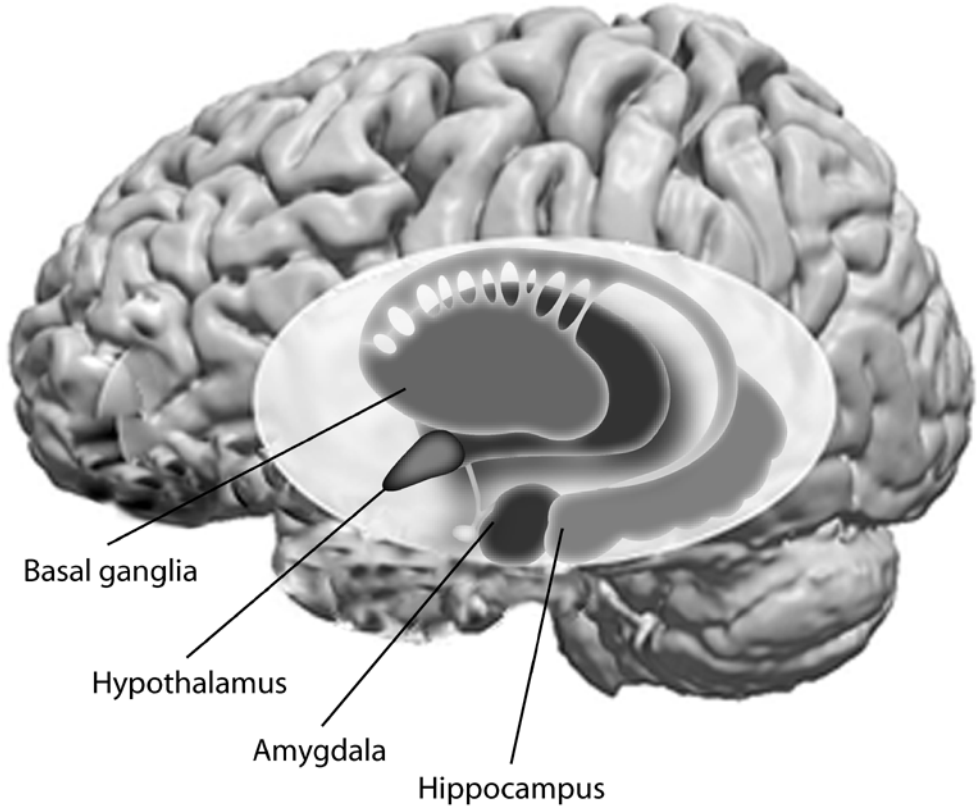


Figure 4.3 Subcortical brain regions important for music. These regions are superimposed on the same scan of author PQP's left hemisphere used in Figure 4.1 and 4.2, for reference.

follows from the way in which neural structures may constrict the flow of water. When a neuron's axon is surrounded by a myelin sheath (thus promoting faster transmission of the neural impulse), the flow of fluid is restricted and will flow in a specific direction. Myelinated axons appear when the brain needs to relay information quickly from one neuron to another. Thus, constraints on diffusion provide a measure of where connections exist in the brain.

Although this technique is relatively new (emerging in music neuroscience research after the year 2000), several recent studies suggest that connectivity between auditory and motor areas, mentioned earlier, may be influenced by musical training, and may be disrupted in people who have trouble matching pitch while singing (see chapter 11). The musicians in our example would probably have large and highly directional connections between these areas.

Methods of cognitive neuroscience

Having surveyed some of the most critical brain regions for music, we now discuss the techniques that scientists use for understanding the structure and function of the brain. Until recently, the only way to understand brain functions (i.e., how it works) was to infer them based on patterns of behavior (as in the lesion/deficit method, described later in this chapter). Now, we have unprecedented tools that allow us to visualize brain activity. Even so, the highly complex nature of the brain's structure and function make it difficult to fully understand what these visual representations mean. Here we discuss the primary methods that cognitive neuroscientists use, and discuss some critical results from research concerning brain processes underlying music and language perception.

To understand how far we have come, it is worth noting that many people used to conceptualize the brain's role in an entirely different way than we do now. In the ancient world, opinion was divided between those such as Hippocrates (460–370 BC), who held that all cognitive functions and all kinds of sensations and feelings originated in the brain, and those such as Aristotle, who credited the heart and liver with at least some of these. By the end of the 18th century, the dominant role of the brain as the instrument for thought, experience, and action was generally accepted. Today there is little doubt that Hippocrates got it right. Nevertheless, one can sympathize with the confusion that existed among ancient thinkers, in an era before medical science had emerged and brain imaging techniques were beyond the scope of imagination.

Thus, with the benefit of technology, we now have an array of techniques that provide different windows into the brain. Even so, the first technique we discuss owes a great deal to thinkers dating far back into history.

The lesion/deficit method

If you know someone who has suffered a serious stroke, you can appreciate the devastating consequences of damage to the brain. Brain damage can render a person speechless, destroy one's ability to walk, or rid someone of the ability to play the piano. Although such consequences are horrible to experience, in a scientific context they can offer useful information that tell us about how the brain works, and even help neurologists treat similar deficits in the future. These observations form the basis for the oldest neuroscience method of researching brain functions: the ***lesion/deficit method***.

Even today, with amazing new techniques available to the researcher, this method still has a good deal to offer. It was first developed during the heyday of the Roman Empire. Galen (AD 129–216), who served for several years as physician to the gladiator school at Pergamum, noticed that gladiators who survived after having received severe head injuries often suffered deleterious changes to their behavior and/or cognitive abilities. Such observations gave rise to studies into the link between brain ***lesions*** (brain damage) and behavioral/cognitive ***deficits*** (loss of function). Here is the logic behind the method: If a psychological function is disrupted by damage to a certain part of the brain that was healthy before damage (or in comparison with healthy brains from similar individuals), one can infer that the region or structure is necessary for that function. Brain lesions are sometimes referred to as 'experiments of nature,' given the cause–effect relationship that one can use to infer the function of affected regions.

Based on this logic, neurologists investigate changes suffered by an individual following injury, stroke, or neurosurgery necessary to preserve one's life (as in extreme cases of epilepsy). Specific lesion sites were once identified postmortem, but now can be viewed using anatomical images of the brain. Advancements in these studies led to the field now commonly known as 'cognitive neuropsychology,' in which changes after brain damage are used to make inferences about the neural mechanisms underlying thought. Such an approach is often associated with the study of single cases, although over time these cases accumulate and allow greater generalization of brain–behavior associations. Persons suffering brain damage (individuals or groups) are conceptualized as a 'treatment group' (the treatment in this case being brain damage!), and are compared to a control group comprising persons with similar backgrounds and of a similar age. For instance, the study of human memory benefited greatly from the observations of how neurosurgery (used to prevent epileptic seizures) influenced the ability of the classic case known as HM to form new long-term memories.

Lesions can be used to map the association between brain areas and musical functions by examining how specific the role of a particular area is. Some brain areas are responsible for a broad range of functions (e.g., the medulla oblongata, located in the brain stem, is vital to our very survival), but we learn more about the brain when a certain area is necessary for some behaviors, but not others. Such effects are called ***dissociations***.

To better understand how to interpret dissociations, we present a hypothetical experiment (which we will use also for subsequent methods) in which a participant is asked to detect altered pitches in familiar melodies. Consider a person who has a stroke that damages the auditory cortex of the right hemisphere. Perhaps this individual cannot accomplish the task in our hypothetical experiment. Imagine that this individual can, however, still recognize changes to the lyrics of these melodies. This hypothetical scenario describes a ***single dissociation*** with respect to the role of the right auditory cortex. The brain area that was damaged seems necessary for recognizing the music-specific features (musical pitch), but not linguistic features (words).

Single dissociations are useful, but limited in the inferences they yield about how different brain structures interact. In this example, there is evidence for music specificity, but we lack evidence for language specificity. Thus it would still be possible that any brain area responsible for language is also necessary for tune recognition. What we need to establish full separation between music and language is called a ***double dissociation***. A double dissociation requires at least two complementary cases. In our example, the second case who establishes double dissociation would have brain damage in a different region from the first (e.g., the left auditory cortex), leading to the inability to recognize changes to song lyrics (language) while still able to recognize changes to pitch (music).

Research using the lesion/deficit method

With respect to the relationship between music and language, case studies of persons with brain damage typically support the view that music and language are neutrally separate, similar to the hypothetical example given above. A number of lesion cases (primarily stroke victims) have been documented in which brain damage hinders language ability (which is referred to as aphasia) while sparing musical abilities (summarized by Marin & Perry, 1999). A smaller number of lesion cases have been found in which brain damage hinders musical abilities while sparing language. For instance, Isabelle Peretz (2003) points to the historic case of Russian composer Shebalin, who suffered a series of strokes in the left hemisphere. Although unable to speak or comprehend language, his musical skills (including composition) were left intact.

In contrast, Peretz documented the case of IR, who suffered bilateral brain damage. This resulted in lasting deficits in her perception and memory of music including the inability to recognize tunes that were familiar to her before the surgery. IR thus acquired a form of ***amusia*** (a term referring to any neurological deficit associated with music), although her language skills remained intact. This is an example of a double dissociation, and has been used as part of an argument that music and language may be represented as separate neural modules (Peretz & Coltheart, 2003).

The lesion method is powerful in that it provides direct causal evidence for a brain–behavior association. However, there are shortcomings to the lesion method. Lesions, whether from accidents of nature or necessary surgery, are not designed to answer a scientific question. Thus, there is no such thing as full experimental control when it comes to such lesions. There are ways to perform studies with greater control. One involves the use of nonhuman animals. However, it is not clear for a domain like music, which is a human creation, how much can be learned from animal models (see chapter 15 for further discussion). The other way involves the use of ***transcranial magnetic stimulation (TMS)***, in which a localized and intense magnetic field can create a temporary 'virtual lesion' without any damage to the brain. Specifically, a virtual lesion can arise when TMS is pulsed at a relatively slow rate (e.g., 1 pulse per second), which leads to neural inhibition. For instance, this kind of TMS disrupts one's ability to synchronize movements to a rhythm, when applied to the junction between the temporal and parietal lobes (Malcolm, Lavine, Kenyon, Massie, & Thaut, 2008).

The lesion/deficit method places the focus on whether a certain brain area is necessary for a certain behavior. Partly due to this focus on discrete structure/function relationships, studies using this technique are often associated with modular as opposed to connectionist view on how the brain works. For instance, the most influential modular theory of music and language is based in large part on studies of lesions (Peretz & Coltheart, 2003). We turn now to a measure of neural activity that is not particularly well suited to questions concerning the necessity of brain areas, but provides an excellent measure of how resource-intensive a particular task is for the brain.

Electroencephalography

For all its merits, there are some considerable limitations to the lesion/deficit method. First, because this method typically relies on isolated cases, it is difficult to know how well the function of brain areas for one case may generalize to the broader population. Even in studies that include multiple similar cases, the specific areas for lesions are never precisely the same, as noted above. Second, the lesion/deficit method doesn't tell us about ongoing brain activity during a particular task. ***Brain activity*** refers broadly to the presence of some kind of effort in a certain area of the brain. In principle, brain activity comes from changes in the firing pattern of neurons, although as we will see, not all measures of activity relate directly to these patterns.

The first technique used to track brain activity was ***electroencephalography (EEG)*** (as shown on the left panel of Figure 4.4a). EEG measures changes in electrical activity by placing electrodes across the entire scalp (as opposed to depth electrodes, which are inserted into the scalp and are rarely used). The transmission of impulses within neurons leads to fluctuations in the electric field surrounding neurons, as positive charged ions flow in and out of specialized gates in the neuron. This leads to a wave pattern associated with electrical activity, not unlike the structure of soundwaves described in chapter 2. Because the EEG measures changing activity over time, it is particularly well suited to explore the *dynamics* of the brain – that is, the timecourse over which your brain responds to events in the environment. As such, EEG is an excellent way to explore domains like music and language in which timing plays such a vital role. For instance, this technique can let us know how long it takes the brain to process the fact that one pitch in a familiar song has been altered.

In addressing this kind of question, researchers typically focus on how the EEG signal changes during a particular mental process. For example, note that at one point the overall pattern of brain activity in Figure 4.4a shifts in the positive direction. This transition could occur while listening to a melody, just after one note goes noticeably out of tune. Such changes in activity are called ***event-related potentials (ERPs)*** because a change in electrical potential results from something a person perceives or does. As Figure 4.4a indicates, ERPs come out more clearly when you average across many trials, which yields a smoothed pattern

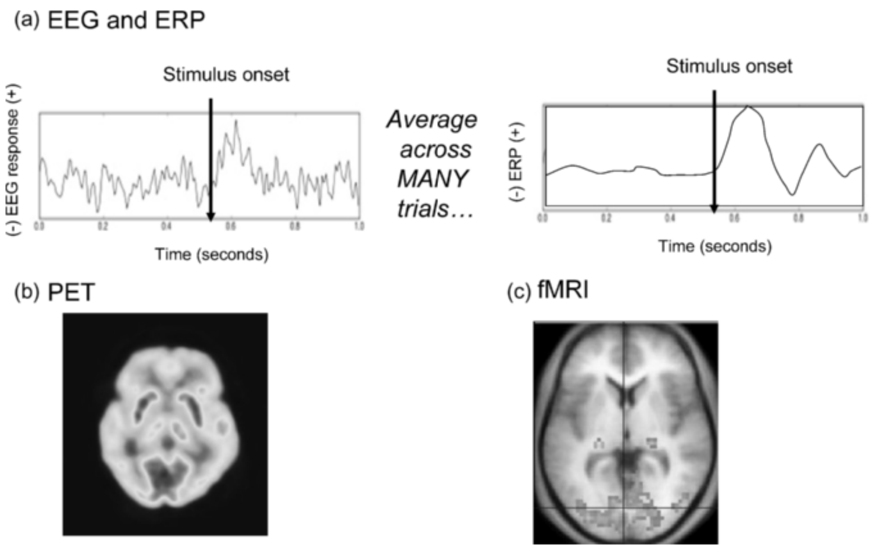


Figure 4.4 Examples of outputs generated by three leading brain imaging techniques. Figures 4.4b and 4.4c show the brain from the top.

Sources: Figure 4.4a: https://en.wikipedia.org/wiki/Electroencephalography. Figure 4.4b: https://en.wikipedia.org/wiki/Positron_emission_tomography. Figure 4.4c: https://en.wikipedia.org/wiki/Functional_magnetic_resonance_imaging.

of change in brain activity over time. The process of averaging, in theory, should eliminate all of the little changes in electrical signals that reflect the steady-state EEG.

Once we have derived an ERP from the EEG signal, the next task is to determine what the ERP means with respect to underlying cognition. This is the tricky part. Roughly speaking, peaks and valleys in ERPs that occur early (e.g., in the first 100 milliseconds) reflect the brain's initial coding of sensory input, whereas later peaks and valleys reflect deeper processing, as in the example we turn to next. For instance, in our example, changes of melody pitches to other pitches that match the musical key may take some time for the brain to process, whereas more obvious changes may take less time. Even though research labs analyze these potentials after averaging across many controlled examples of a phenomenon (e.g., very many trials including either no pitch change or an unexpectedly mistuned pitch), the underlying neural phenomenon is a general one. For instance, if you hear a recording of the famously off-key singer Florence Foster Jenkins (dramatized in a 2016 movie by the actress Meryl Streep), your brain will probably generate a pronounced ERP response when the high pitch in Mozart's 'Queen of the Night' aria (from the *Magic Flute*) is sung noticeably flat.

EEG/ERP research on the music–language link

Some of the earliest evidence for the integration of music and language stemmed from research using ERP technology. In speech, if the ordering of words in a sentence (i.e., the syntax) is unusual or even incorrect, there is an increased positive polarity about 600 ms after the point at which the syntax breaks down (Osterhout & Holcomb, 1992). For instance, 'The broker persuaded to sell the stock was sent to jail' is grammatical, but at the point when one hears the word 'to,' the sentence sounds ungrammatical because one expects an object to the verb 'persuaded.' Note that there is no significance to the negative or positive direction of the ERP;

only the timing and the amplitude (magnitude of change) matter. Many have proposed that music also is guided by syntactic rules, roughly analogous to those in language (e.g., Lerdahl & Jackendoff, 1983; see chapter 7 in this volume). In keeping with this view, a similar ERP response to music happens when listeners hear an unexpected chord change – a violation of musical syntax (Patel, Gibson, Ratner, Besson, & Holcomb, 1998).

Recently, other measures of the EEG signal have been used to measure brain activity at a smaller (i.e., more rapid) time scale than ERPs do. A particularly influential measure in the past several years has been the use of Fourier analysis (see chapter 2) of the EEG waveform and how the pattern of frequencies changes with the stimulus. Amazingly, although these responses are measured at the center of the scalp, such short-latency responses in fact originate in the auditory brainstem (see chapter 3), and thus reflect very early auditory processing.

One such analysis that is related to the association between music and language is the ***frequency following response***. This comes from analyzing the fundamental frequency of the EEG response, just as one would an auditory signal (see chapter 2). Amazingly, when one analyzes the fundamental frequency (F0, also discussed in chapter 2) of the EEG signal while a person hears changes in pitch, the F0 of EEG shifts in a way that literally follows the F0 of the sound! So in a very real sense, brainwaves recreate sound (for a demonstration, see http://www.brainvolts.northwestern.edu/demonstration.php). Physiologically, these rapid responses relate to sound processing in the brainstem (see chapter 3), even though the recording site is an EEG sensor on the skull. But some brains exhibit a more faithful frequency-following response.

Here we have some evidence favoring associations between music and language processing. Wong and colleagues (Wong, Skoe, Russo, Dees, & Kraus, 2007) examined the frequency following response to Mandarin tones. Mandarin, the most common language of mainland China, is a ***tone language***, in which the pitch used to produce a word conveys word meaning (by contrast, in English we use pitch primarily to convey emphasis). For instance, in Mandarin, syllables can be spoken with a rising falling, high stable or low stable pitch. Changing the pitch pattern you use can change a word, so that 'ma' may mean either 'mother' or 'horse.' Among a sample of English-speaking participants, those with musical training exhibited a more accurate frequency following response to Mandarin tones than those without training. As such, musical training may attune our brain's sensitivity to spoken pitch, even for languages we do not (yet) understand.

ERPs have excellent ***temporal resolution***: They measure the timing of neural responses with high precision (within milliseconds). Thus, ERPs are exquisite measures of neural dynamics, such as the surprise one's brain exhibits after hearing that a familiar melody's pitch has been changed. On the other hand, ERPs are not good at locating the precise brain region associated with a process (i.e., their ***spatial resolution*** is poor) because brain activity is recorded at the scalp, whereas the underlying activity begins at the synapse and spreads throughout the brain. For this reason, Patel (2008) characterizes ERP as a good way to measure the use of neural resources (i.e., neural effort), while not focusing on neural representations (i.e., brain areas) that use those resources. The techniques we describe next offer much better estimates for the location of brain activity, as a way of addressing neural representation.

Positron emission tomography and magnetic resonance imaging

In this section we describe brain-imaging techniques designed to pinpoint the location of brain activity. These techniques work on the premise that brain regions require more energy-rich glycogen, and more oxygen necessary to metabolize it, when they are engaged in a particular activity. These techniques also yield the images that are often described as showing the brain 'lighting up,' as in the popular press article cited at the beginning of this chapter, although not really accurately capturing the brain's response. There are two such techniques; the older one is called ***positron emission tomography (PET)*** (illustrated in Figure 4.4b).

Whereas in an ERP study, one is fitted with a cap containing many electrodes, the following techniques involve much more elaborate equipment. In these experiments, you would lie supine with your head inserted into a large tube or 'bore' that contains specialized sensors. This can lead to an intimidating experience for the participant, a fact that must be dealt with by those conducting the study. But there is a decisive advantage to these other techniques: They can pinpoint what areas of the brain are involved in a certain task, including what areas are involved in recognizing pitch changes to a familiar tune.

It is critical to note that for both of these techniques, the images of brain activity usually do not show activity during a single task, but instead relate to differences in activity across tasks. Traditionally, researchers use the ***subtraction method***, and display differences in activity between an experimental task (e.g., listening to a familiar song with altered pitches) and a control task (e.g., or listening to a familiar song with unaltered pitches, or just resting). This is one reason why one should not take too seriously claims about a given task involving the 'whole brain.' One can observe a lot of brain activity even when the brain is in a resting state, including several whose activity constitute the brain's 'default mode' network (Raichle et al., 2001).

Positron emission tomography

In the PET procedure, an atom of a radioactive isotope of a common element, usually fluorine, is attached to a molecule of glucose, the blood sugar that provides the fuel for brain activity. This material is injected into the blood stream of the participant. Thus, if you participate in a PET study, you would do so with a needle inserted into your arm, similar to what one might experience during a hospital stay.

The purpose of this invasive procedure is to obtain a direct measure of blood flow. During brain activity, there is an uptake of glucose, and with it some of the molecules seeded with the radioactive fluorine. As the radioactive fluorine atoms decay, they emit a positron (a positively charged electron). The positron soon meets a negatively charged electron, and the two annihilate each other, emitting a gamma ray. Increased blood flow leads to more frequent collisions. The aforementioned sensors that surround the participant's head record the rays that result from these collisions. In this way, PET offers a direct and highly sensitive measure of blood flow within the brain. With respect to our example, PET could inform us as to the areas involved in detecting pitch changes, with greater precision than either the lesion/deficit method or using ERPs. At the same time, PET images are a bit blurry because the positions of gamma rays do not precisely identify the location of collisions.

The use of PET has led to several important discoveries. For instance, the study of ***auditory imagery*** for melodies (the conscious mental simulation of auditory perception) was difficult to carry out with behavioral methods because mental images are by definition subjective and internal. However, an influential study of auditory imagery with PET showed that auditory imagery relies on many of the same brain regions as auditory perception, with the significant exception of the primary auditory cortex (Zatorre, Halpern, Perry, Meyer, & Evans, 1995). They asked participants in the PET scanner to think of a familiar melody (e.g., 'Jingle Bells') and imagine if the pitch associated with one word (e.g., 'snow') is higher or lower than the pitch associated with a different word (e.g., 'sleigh') in the song lyrics. In this case, as with many others, PET helped us better understand the nature of mental representations underlying a nebulous but important aspect of musical experience. However, because of the slight blurriness of PET images and the moderate exposure to radiation that the technique involves, a different technique has dominated neuroimaging research more recently.

Functional magnetic resonance imaging

The technique known as ***magnetic resonance imaging (MRI)*** provides clear pictures of brain anatomy, and the related use of ***functional magnetic resonance imaging (fMRI)*** provides precise estimates of the location of brain activity (i.e., function).

We start with MRI, a fairly common procedure that you may have experienced at a hospital. An anatomical MRI measures tissue density of internal organs with incredibly high precision, so that a doctor may see clearly where aberrations in brain structure exist. Like TMS, MRI relies on the use of magnetic fields, although with MRI, the magnetic field is not used to disrupt neural activity, but rather to disrupt the atomic spin within neural tissue. An intense magnetic field (only harmful to your credit card if you are foolish enough to keep it with you) is used to align the magnetic fields of individual hydrogen atoms, whose magnetic fields are typically unrelated. Then a radiofrequency pulse is sent through the brain to perturb the alignment of these atoms. When the pulse ends, the atoms return to the alignment that is brought about by the magnetic field, and the MRI uses information about how many atoms are changing position to measure tissue density. In clinical contexts, MRI helps a doctor identify what areas of a brain may be damaged or malformed. For fMRI studies, however, the MRI provides a map of where brain activity may be happening.

When you look at images of brain activity in a journal or magazine, you see MRI images with pictures of brain activity superimposed on them. These colorful images of brain activity come from fMRI (see Figure 4.4c). Like PET, fMRI uses blood flow as an indirect measure of brain activation, but unlike PET, fMRI does not directly measure blood flow. When a region of the brain is active and thus uses more blood, there is a commensurate increase in the use of oxygen by the blood. This leads to a change in the proportion of oxygenated and deoxygenated hemoglobin in that region. Deoxygenated hemoglobin is more sensitive to magnetic fields than oxygenated hemoglobin; fMRI uses this differential sensitivity to generate a measure of the so-called ***BOLD (Blood Oxygenation Level Dependent) response***. The fMRI image is blurry, whereas the MRI image is impressively detailed; therefore researchers usually use both images and superimpose the fMRI image on an MRI image. Because fMRI does not use a tracer to measure blood flow directly, you would not need to have a needle in your arm while participating in an fMRI study.

As might be expected, most fMRI studies have employed adult participants; an example of a participant taking part in an fMRI study of music performance (and therefore playing a keyboard) is shown in Figure 4.5. However, researchers have recently begun to modify these procedures for use in research with child participants. For instance, Overy and colleagues adapted the procedures by preparing child participants with simple task instructions ('be a listening detective'), providing a cartoon story of a child having an fMRI scan, allowing children to take a soft toy into the scanner with them, shortening scanning runs, and using techniques that reduce the loud fMRI scanner noises (Overy, Norton, Cronin, Winner, & Schlaug, 2005).

PET and fMRI research on the music–language link

What about the neural localization of music versus language? A generalization in the literature is that PET and fMRI studies most often show overlap between music and language,

Figure 4.5 Experimental participant with keyboard, preparing for an fMRI experiment concerning music performance. Prior to the experimental session the participant will be moved back so that her head sits within the bore of the scanner.

Source: Photograph used with permission from the participant, and the Buffalo Neuroimaging Analysis Center. Copyright © Peter Pfordresher.

in contrast to the lesion/deficit method (Peretz, 2012). This is because the lesion/deficit method, as discussed earlier, focuses on whether an area is necessary. By contrast, PET and fMRI reveal many areas that may form an integrated network, but that have varying roles – some necessary, others not. Although this generalization works in many cases, as we will see, there are diverse results from studies using PET and fMRI.

In fact, one of the classic early PET studies supports the modular view of music and language. This study considered the way in which the brain processes speech when one attends to pitch qualities of speech, similar to when we listen to music (Zatorre, Evans, Meyer, & Gjedde, 1992). More frontal areas of the brain became active when people focused on pitch qualities of speech, as opposed to when they listened without any fixed intention ('passive' listening). Activity in these frontal areas likely reflected conscious processing of meaning. Furthermore, when people listened to speech and attended to phonetic aspects (i.e., the smallest units of meaning, analogous to spoken letters), brain activations were more left-lateralized, whereas attending to pitch made brain activations more right-lateralized. Thus, this study supported the view that the left hemisphere plays a dominant role in speech perception, whereas the right hemisphere is better at processing music, or 'music-like' qualities of speech.

A more recent fMRI study also supports the idea that music and language are processed in different hemispheres (Wong, Parsons, Martinez, & Diehl, 2004). This study focused on the perception of lexical tones in tone languages, discussed earlier in the section on EEG. Listeners who either spoke Mandarin (as well as English) or spoke only English heard pitch contours that could be either English or Mandarin. English monolinguals processed both contours in the right insular cortex (deep within the brain, at the juncture of the temporal, parietal, and frontal lobes; see Figure 4.1). Likewise, when bilingual Mandarin/English speakers heard the pitch contour in an English context, activations appeared in the right insular cortex. However, when bilingual Mandarin/English heard the pitch contour in a Mandarin context, brain activations switched from the right insula to its left homologue (i.e., the same structure, but on the left side). The implication here is that speakers of tone languages treat pitch as functionally different within that tone language, whereas pitch within languages like English has a more 'music-like' role, and is processed accordingly. However, you need to know a tone language in order to process its pitch in the brain's 'language' area – so the distinction is not in the stimulus, but in the mental processing of that stimulus.

It is important to note that the aforementioned lateralization effects are not of an 'all or none' variety. Although music and language were differently lateralized, in each study there was some activation in both hemispheres for both conditions. As such, other research has addressed what areas within a hemisphere are used to process music and language. This perspective has yielded some support for the notion that music and language are integrated in the brain. Consider the brain regions around Broca's area (ventral part of the left frontal lobe), which was long considered to be responsible for controlling speech production. fMRI research now suggests that this part of the brain also responds to musical structure. For instance, the presence of unexpected chords in music stimulates activity in Broca's area (Koelsch et al., 2002). In another study, people heard normal or scrambled versions of classical music that was either familiar or unfamiliar to participants (Levitin & Menon, 2003). Listening to normal music, as opposed to scrambled variants, elicited activation in the pars orbitalis region (see Figure 4.2), a region near Broca's area also traditionally associated with language processing. These findings support a newer view of Broca's area and surrounding regions as being responsible for processing syntax, thus complementing the aforementioned study of Patel and colleagues (1998), which suggested that the brain responds similarly to deviations of syntax for music and language.

The research on music and language in the present chapter provides a good example of what neuroscience has to offer. For years, people have pondered whether 'speech is special' (Galantucci, Fowler, & Turvey, 2006). The neuroscience research so far suggests that this kind of assumption is too simplistic. But why should the answer be simple? Music and language are complex, dynamic forms of expression that people use and interpret in differing ways.

Furthermore, the brain is an organ we are getting to know, but are far from fully understanding. One weakness of traditional fMRI analysis, which focuses on activations within specific areas, is that it is hard to identify when areas are working in networks. Recent techniques have been designed to help us better understand temporal patterns of neural activation, and how activation in certain areas may lead to activations in other areas. Recent studies that have adopted these techniques suggest that overlapping activations for music and language may reflect different temporal activation patterns within these areas (Norman-Haignere, Kanwisher, & McDermott, 2015; Peretz, Vuvan, Lagrois, & Armony, 2015). Finally, as we consider in the next section, our brains are not static in structure and function, but are constantly changing and developing.

Neural plasticity and individual differences

It is important to note that associations between brain areas and their functions are not fixed. So far, we have focused on brain structure and function in the healthy adult brain. Even here there is variability, and when one looks at brains developmentally and across special populations (including highly skilled musicians, as well as stroke victims), differences are magnified. This kind of variability is important, and of keen interest to cognitive neuroscientists. In fact, one reason why music training and performance is of interest to neuroscientists is because of their potential to demonstrate ***neural plasticity***, our nervous system's capacity for change.

Some time ago, the neurosurgeon Wilder Penfield (1891–1976) explored the brains of people who were about to have surgery for epilepsy. In order to determine what aspect of function could be damaged by this surgery, he stimulated different parts of the brain. This led to the discovery that regions of the motor cortex (see Figure 4.2) had a direct relationship to muscles in the body, and that the spatial mapping of the brain had some relationship to the spatial mapping of the body. For instance, neurons in the motor cortex's 'hand region' are laid out in a way that forms a hand-like shape. Furthermore, the amount of space devoted to these representations related to the degree of precision in motor control: for instance, more space is devoted to the hands than to the feet. Given these results, one might expect that musical training may influence the neural representations devoted to the hand, or some other effector used to perform music.

In fact, musical training shapes the fine structure of the brain's hand region. For instance, the area responsible for controlling the left hand (which is in the right hemisphere) responds more strongly to tactile stimuli among violinists (who use that hand for fingering) than among those without such training (Elbert, Pantev, Wienbruch, Rockstroh, & Taub, 1995). In another study, it was found that anatomical asymmetries in the brain varied with both the extent and type of musical training (Bangert & Schlaug, 2006) in a way that reflected the use of left versus right hands for pianists (more right-hand representation in the brain, as most piano music requires more dexterity in the 'melody' hand than left-hand accompaniment) versus violinists (more left-hand representation, as playing violin requires more movements of the fingers of the left hand than the right-hand, which holds the bow). The effects of music training and the neural mechanisms involved in executing a music performance will be discussed further in chapters 10 and 11.

Coda

It is astonishing how much more we know about the brain now than we did only a few decades ago, as exemplified by how technology has furthered our understanding of the neural basis for music and language. Neuroscience is well equipped to answer questions regarding the neural processes that support the perception of sound as music. It is important to realize, though, that most of the underlying theoretical issues originated in research that relied on measurements of overt behavior (responses on a questionnaire, performance on a task, etc.), not neural activity. Beyond this, the notion of what neural 'activity' constitutes is not clear-cut, given that the most common measure – in fMRI – is in fact a measure that is twice removed from neural action potentials (i.e., the BOLD response that is associated with blood flow that is further associated with action potentials). Thus our vision of the future is not one in which neuroscience eclipses behavioral research into the conditions for musical experiences, but rather it is one in which the study of brain processes and the experiential and performance aspects of musical life complement each other.

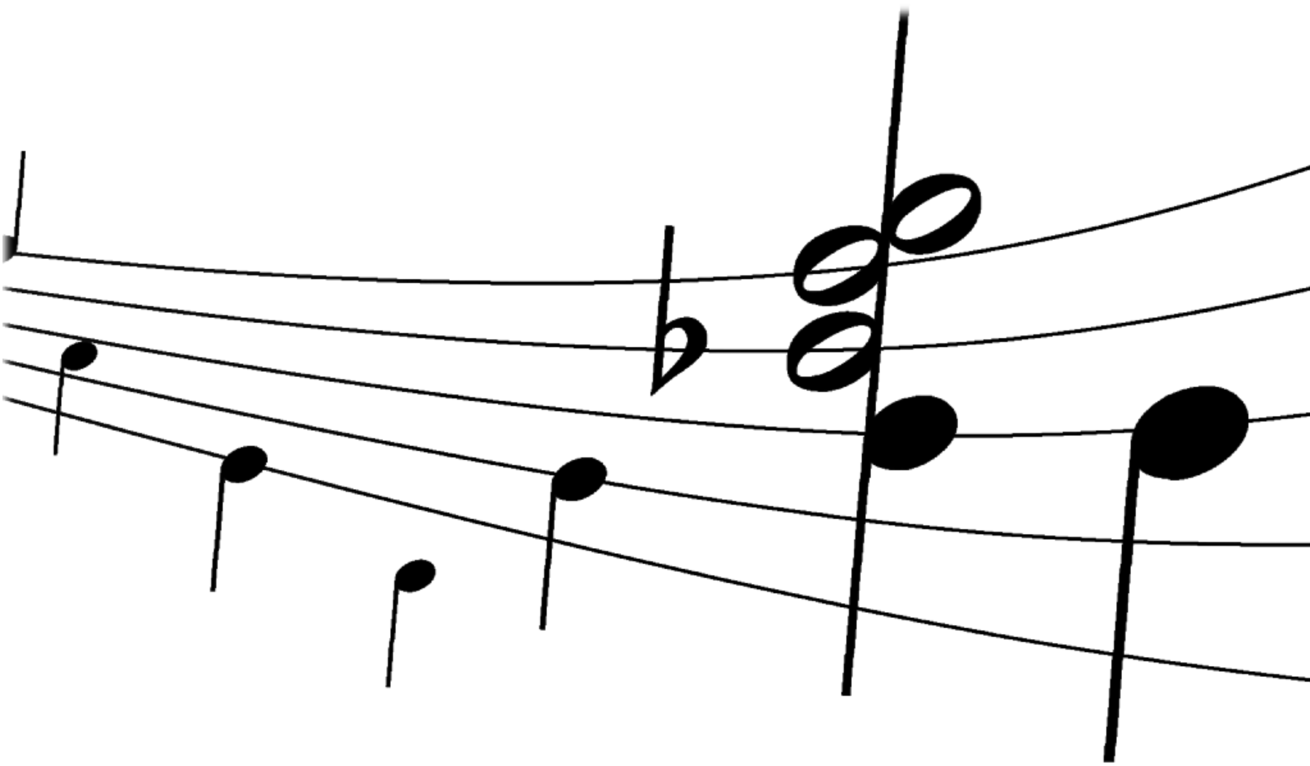
In this context, it is important to recognize how much we do *not* know from reading brain images or researching lesioned patients. What is the origin of the creative 'spark' that leads to a beautiful composition? Or a particularly moving or expressive performance? Is it neural activity? If so, what initiates such activity? These are deep issues, and the answers will probably not be answered by localizing brain regions correlated with such complex and multifaceted activities as music composition or performance (further discussed in chapter 11). At a public lecture during the 2015 meeting of the Society for Music Perception and Cognition, the neuroscientist Charles Limb reflected that neuroimaging techniques do provide a window into the brain … but the view we get is similar to peering into 'a small window in the basement of someone's house.' That is, we get a clearer view of neural functioning than was once imaginable, but we still see only a small part of the brain's activity.

Note

1 Whereas language refers to a general communication system that can be conveyed through the voice, writing, or gesture, speech refers specifically to the voice.

Part II

The perception and cognition of music



5 Perception of musical pitch and melody

The following case of a 61-year-old Canadian man was documented by Peretz (1993):

> GL was referred to us in 1989 because of his persistent amusia . . . . GL stated that he was totally unable to pick out familiar music and did not enjoy music any more. His musical background is that of a nonmusician who was nonetheless an avid listener to popular and classical music, attending concerts and musical recitals very regularly before his illness . . . . Contrasting with his relatively good linguistic abilities, recognition of tunes (without lyrics) was totally abolished. Out of 140 musical excerpts (including his national anthem) that are very familiar to everybody in Quebec, he could not identify a single one.
>
> (p. 27)

About 10 years earlier, the man had suffered an aneurysm on the right side of the brain. The year after that, he suffered a mirror aneurysm on the left side, after which he was diagnosed with severe ***Wernicke's aphasia*** and also ***amusia*** (broadly defined as loss of language comprehension and music ability, respectively). Although he recovered most of his speech abilities, the amusia persisted. When tested on his ability to discriminate isolated pitches, GL's performance was comparable to matched controls. However, clear deficits emerged when pitches were presented in a melodic context. For instance, GL had difficulty identifying whether two melodies that differed by one pitch were the 'same' or 'different.' He was also unable to correctly identify whether a melody sounded 'complete' (ending on the tonic) or 'incomplete' (not ending on the tonic). As described above, he could not identify well-known melodies such as *O Canada*, his own national anthem! Peretz concluded that GL could comprehend isolated pitches, but had 'lost access to tonal knowledge' (1993, p. 51),[1] which is absolutely critical if one is to 'make sense' of melody.

The case of GL shows that we should not take the perception of melody for granted. If melodies are made up of separate pitches, what binds a melody together so that it is heard as an organized and coherent tune? Melodies are generated by arranging a small set of discrete pitches and durations into an infinite number of possible patterns. How are we able to remember specific arrangements? That is, how do we learn and retain the thousands of tunes we 'know'? These questions serve as the starting point for our discussion on melody and pitch.

The elements of melody

In keeping with existing norms (e.g., Radocy & Boyle, 2003), we define a ***melody*** as a sequence of pitches that sound like they belong together. This rather open-ended definition captures the idea that the meaningfulness of a melody lies in the global pattern formed by pitches – that the 'whole' is more important than the individual 'parts.' While each tone of a melody reaches the listeners' ears as if it were a single bead, listeners 'thread' the beads together into strands. At the outset of this chapter, we will explore some essential elements of melody that help us understand how melodies are formed: pitch, interval, contour, harmony, and key.

Pitch is more than high versus low

Pitches are the basic units of melody. In chapters 2 and 3, we discussed pitch in a familiar way: as a perceptual quality of sound that ranges from low to high. However, as we have seen in the case of GL, our perception of pitch is more complex within the context of a melody. Specifically, in music our experience of pitch becomes multidimensional and nonlinear.

How can we say that a percept that is related to a single physical dimension (frequency) can be multidimensional? When we speak of a psychological dimension, we refer to some scale on which different perceptual experiences can be distinguished. Remember that perceptual experiences do not have to replicate the physical stimulus exactly. In vision, it is possible to see things that aren't part of the physical stimulus (e.g., illusory contours), or fail to see things that are there (ultraviolet light). The visual perception of color, however, provides the best example of a multidimensional perceptual experience. Color is related to a single physical dimension: wavelength. However, the colors we experience can be distinguished by three different perceptual dimensions: hue (the color's category), saturation (how pure or vivid the color is), and brightness.

So it is with musical pitch. Roger Shepard (1982) argued that musical pitches may be based on as many as five dimensions. We focus on two here. One dimension, true of pitch either in or out of a musical context, is ***pitch height***, which is related to the frequency of vibration. However, a more musically important dimension of pitch is the second dimension, ***pitch chroma***. Chroma refers to the category (or 'class') represented by a certain pitch. The names we give notes in Western tonal music (e.g., C, D, E) refer to pitch chromas. Chromas are identical when separated by an ***octave***, a 2:1 ratio of a tone's fundamental frequency, and thus constitute a distinct dimension from pitch height. The fact that tones separated by an octave inhabit the same chroma category is called ***octave equivalence***. An important implication of octave equivalence is that melodies played an octave apart may be heard as identical. This helps, for instance, when men and women sing in unison. Most female voices are best suited to pitches about an octave higher than most male voices. Yet men and women often sing the same 'notes' together. This could not happen without octave equivalence.

Octave equivalence changes everything. Because of octave equivalence, the experience of similarity in pitch cannot be linear; it becomes cyclical. Consider an ascending major scale (*do-re-mi...*). As you go up this scale, the pitches change and become higher and higher. But do they become increasingly dissimilar from the starting pitch (*do*)? Perhaps for a while, but then at some point you return 'back to *do*' (just as in the famous song from *The Sound of Music*). Because the sequence of chromas that you sing repeats itself every octave, relationships among chromas are said to form a circle rather than a straight line. This leads to the ***chroma circle***, a representation of how the listener perceives chromas. In fact, the name chroma is based on the fact that this circular arrangement of pitch resembles the way we perceive relationships among colors based on their hue (the color wheel). The arrangement of chromas in a circle is represented as the base of the helix in Figure 5.1.

Granted the critical importance of pitch chroma, pitch height still matters in music. Consider pitches separated by an octave. Although they share the same chroma, listeners can

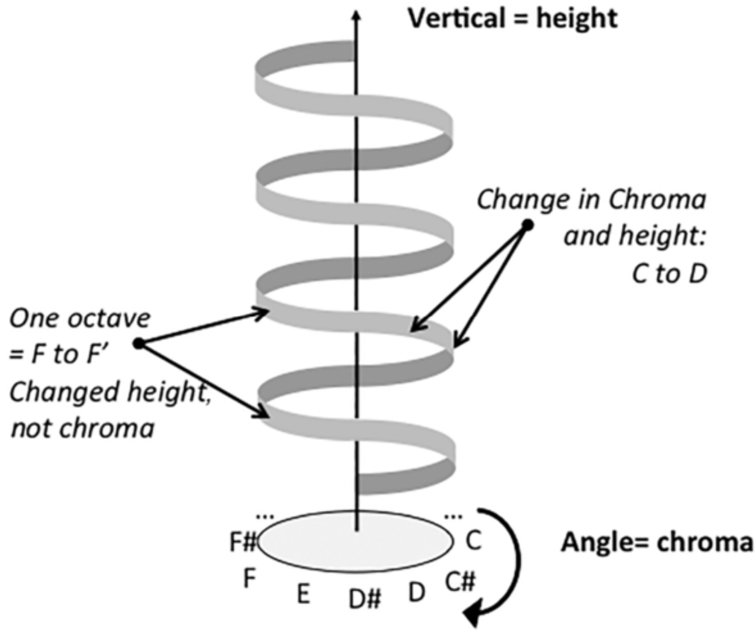


Figure 5.1 The simple pitch helix. The ribbon shape represents pitch relationships that vary in height (vertical dimension) and chroma (angle implied by use of perspective). The chroma circle is shown for reference at the base of the figure, and the arrow pointing upwards reflects variations in pitch height.

Source: Adapted from Shepard (1965, Figure 8, p. 105) with permission from Stanford University Press.

easily distinguish these pitches. Musical pitches are distinguishable both by how high or low they are in frequency (as in the psychoacoustic definition, chapter 3), and by their chroma. The way these two dimensions interact is represented by the ***pitch helix***, shown in Figure 5.1. This is a shape that represents how we hear relationships among pitches based on chroma and height. Any perceived pitch is thought to fall somewhere on the spiral shape of the helix. Pitches that fall near each other are heard as more similar to each other, and pitches farther apart are heard as more distinct. Notice that every angular turn around the circular dimension (chroma), results in a change in vertical height. This characteristic relates to an important aspect of perception: Every change in chroma necessarily involves a change in pitch height. However, the converse is not true. It is possible to change height without changing chroma. On the helix, this involves a shift in the vertical dimension that remains on the same angle around the circle. Musically, this is simply a change in octave.

Next, we consider how listeners categorize pitches in perception, and encode pitch information in memory. There are two ways in which this can happen. One can categorize and encode pitches based on their individual properties (***absolute pitch***), or their relationship to a larger context (***relative pitch***).

Absolute ('perfect') pitch: Rare gift or common ability?

Every chroma in Western music has a specific label, often represented as a letter (A through G) that represents a spot on the musical staff as well as a white key on the piano. This letter

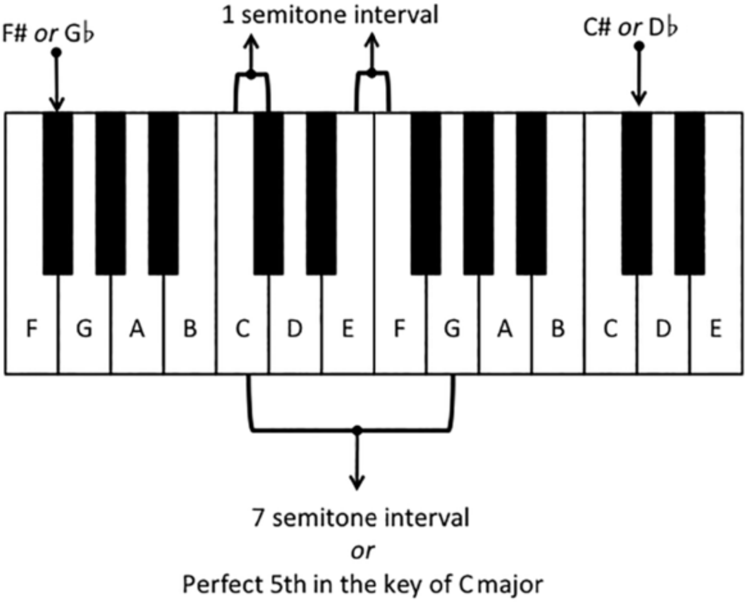


Figure 5.2 Illustration of relative and absolute pitch categories on a piano keyboard. Absolute pitch categories are associated with letter labels for single keys (white or black), whereas relative pitch is based on relationships between keys in semitones or based on scale steps, as in the perfect fifth.

Source: Copyright © Peter Pfordresher.

may be accompanied by an 'accidental,' which is a symbol indicating that the pitch should be shifted to an adjacent black key on the piano that is either higher (sharp, ♯) or lower (flat, ♭) than the referent pitch. Figure 5.2 illustrates the relationship between selected pitch labels and piano keys (higher pitches are to the right). This figure also illustrates an important unit of measurement for pitch differences: the semitone. A ***semitone*** is the smallest possible unit of change between any two pitches in the Western system; it is the difference between any two adjacent piano keys, which can be two white keys (e.g., E and F; see Figure 5.2) or a white and black key (e.g., F and F♯). Finally, one may see some indication of the octave in which the pitch occurs: C1 is an octave below C2, or more abstractly C is an octave lower than C′. These labels come from the conventions of music notation.

A more interesting question for psychology involves the relationship between these labels and our experience of music. Labels like C♯3 constitute an ***absolute pitch (AP)*** category, because the pitch in question is defined independently of any surrounding context. Absolute pitch categories are useful for notation and music theorists who describe the structure of music verbally, but do they have any relationship to everyday musical experience?

Labeling the absolute pitch of musical tones

It is commonly believed that most listeners do not encode the absolute pitch of melodies. This belief comes from the fact that the standard traditional way to measure what is called absolute (or 'perfect') pitch ability involves a task in which one has to label individually presented pitches without being provided with a reference pitch. It is rare to find individuals who perform well on this task, with informal estimates ranging from 1 in 1,500 (Profita & Bidder, 1988) to 1 in 10,000 (Bachem, 1955). It should be noted that these estimates (which are cited often) are not based on formal observations, which would be enormously difficult to carry out, and thus should be interpreted cautiously (Miyazaki, Makomaska, & Rakowski, 2012).

Perhaps in part because of its apparent rarity, the origin of the ability to label absolute pitch is a topic of widespread fascination. Moreover, this ability is often associated with astounding levels of musicianship: Mozart was renowned for his ability to perceive the pitch of everyday objects, like a doorbell. People like Mozart are often said to 'have AP,' as if performance on the labeling task is commensurate with the ability to hear AP. As we will see later, such claims are problematic, yet the factors that lead one to 'have AP' are nevertheless fascinating.

What is the source of this uncanny skill? One important component, and possibly the most important, may be starting music lessons at an early age. Given that labeling of pitch is involved, musical training is of course necessary for AP. However, there is a dramatic difference in the frequency of this ability for those who begin lessons under the age of 6 years compared with those who begin lessons later in life (Baharloo, Johnston, Service, Gitschier, & Freimer, 1998; Takeuchi & Hulse, 1993). But why should this occur? One possibility is that musical training allows performers to associate pitches with verbal categories.

Russo, Windell, and Cuddy (2003) tested this possibility directly by having children aged 3–6 years, as well as adults, learn to discriminate a single 'special' note (C5) from a set of other notes that served as distracters. All participants improved on the absolute identification of this note (which is precisely how absolute pitch is typically measured), but improvement was best for children who were 5–6 years old. This study thus offers a model of how absolute pitch might be learned, and also offers direct support for the notion of a 'sensitive period.'

Other studies suggest that the type of training, and not just the age of onset, may matter. In comparing Japanese to Polish conservatory students, Miyazaki and colleagues (2012) found higher rates of AP among Japanese students that was independent of the age at which training began and the intensity of training. The authors argued that a likely contributor was the fact that many Japanese students (importantly, those most likely to have AP) underwent training that used a fixed-*do* solfege system. This system uses the standard *do, re, mi…* labels for pitch, but is set up so that *do* always refers to the same pitch chroma. This contrasts with the movable-*do* system used elsewhere, including Poland, where *do* can refer to any tonic pitch (e.g., C in C major, but D in D major).

Another contributing factor to AP could be language. Deutsch and colleagues (Deutsch, Henthorn, Marvin, & Xu, 2006; Deutsch, Dooley, Henthorn, & Head, 2009) have documented higher rates of AP among Chinese conservatory students who speak an Asian ***tone language*** fluently (a language in which pitch conveys word meaning; see chapter 4 for further discussion). It is important to note, however, that Japanese is not a tone language, and thus it is not clear whether the use of pitch in Japanese was associated with the increase in AP observed by Miyazaki and colleagues.

So, can anybody with the right kind of early experience acquire AP? Probably not. There is evidence that some people may be genetically inclined to acquire this ability. Profita and Bidder (1988) observed that AP traveled in families. Later, Baharloo and colleagues carried out an extensive study to quantify the genetic basis of AP among musical families using a measure called 'sibling recurrence risk' (Baharloo, Service, Risch, Gitschier, & Freimer, 2000). The term 'risk' in this measure comes from the fact that researchers in genetics are usually concerned with the likelihood of inheriting diseases (like schizophrenia) or other impairments such as color-blindness rather than a desirable trait like musical ability. The sibling recurrence risk is a ratio contrasting the probability of exhibiting some phenotype

(AP, schizophrenia, autism, etc.) when your sibling has it, relative to the probability of having that phenotype for anybody in the general population. The authors estimated this ratio for AP to be between 8:1 and 15:1. In other words, the 'risk' of having absolute pitch given that a sibling has it may be 15 times greater than if your sibling is a musician without absolute pitch. By comparison, the ratio for schizophrenia is 9:1, which is high for mental disorders. Another team of researchers performed more detailed genome-wide analyses of absolute pitch possessors and identified specific chromosomes associated with this ability (Theusch, Basu, & Gitschier, 2009)! This is impressive work; but it is important to remember that genes always interact with the environment.

As mentioned in chapter 3, pitch perception changes with age due to physiological changes. As such, it is plausible that AP would change with age. One of your authors (ST) was identified as having AP at around age 7, while being tested by an examiner during an aural exam for the Associated Board of the Royal Schools of Music. She was quick at identifying single pitches, and could even name pitches comprising three- or four-note chords with high accuracy. During her mid-30s, she began to make more errors when identifying pitches (especially in the extreme high and low ranges), and tended to be a semitone sharp when incorrect (e.g., hearing the lowest E♭ on a grand piano as an E).

She has since discovered that this 'shift' is common. Athos and colleagues (2007) tested almost 1000 AP possessors between the ages of 8 and 70, and found that 'pitch errors [among individuals with AP] increase with age, and they tend to be sharp' (p. 14796). In fact, one 44-year-old participant did not identify a single pitch correctly during the test, but was consistently one semitone sharp, a tendency he had observed since age 22. The reason for this shift is not yet known, but Athos and colleagues (p. 14797) surmise that it may correspond with some 'age-dependent physiological change that alters the mechanical properties of the cochlea' (these properties are further discussed in chapter 3). Indeed, the authors hypothesize that 'such a gradual perceptual shift is common to most people as they age, yet they are unaware of it unless they have AP' (p. 14797).

Memory for absolute pitch

It is important, however, to recognize the potential importance of labeling in standard measures of absolute pitch perception. The need to associate pitches with verbal labels automatically eliminates from consideration anybody without formal musical training, and focuses on a very specific musical ability – the formation of fixed associations between pitches and a set of culturally defined labels. When one removes the labeling component of the test, absolute pitch perception may in fact be quite prevalent. In other words, the ability to label pitches absolutely may be rare, but the ability to remember absolute pitch may be common. So as to avoid confusion with the traditional sense of AP as a labeling ability, we will refer to absolute memory for pitch as ***AP memory***.

The first resounding demonstration that AP may be more common than often assumed was reported by Levitin (1994). He simply asked participants to sing their favorite tune from memory, beginning wherever they wished. If people in general do not have AP memory, then we should find that people typically transpose melodies into the most comfortable key within their singing ranges. This would lead to a fairly uniform distribution of produced pitches across a sample of individuals for a given tune. Pop tunes, after all, are not often in a comfortable register, particularly for the male voice. However, participants (who were not trained singers and did not possess the AP labeling ability discussed in the previous section) usually started singing melodies on the original starting pitch (within one or two semitones), even for larynx-strainers like 'Hotel California.' Importantly, pop songs used in Levitin's study were those that listeners had always heard in the same key.

This study showed that although correct *pitch-labeling* (C, D, E, etc.) without a reference pitch may be rare, good *pitch memory* seems to be widespread. Such use of AP memory in singing is not limited to pop tunes. Prior to Levitin's study, Halpern (1989) demonstrated that people tend to sing familiar folk songs, such as 'Happy Birthday to You,' using consistent pitches, even when these songs are not associated with a specific musical key. Similarly, Bergeson and Trehub (2002) recorded the pitches mothers chose to sing to their babies, and found surprising consistency within individuals from one week to the next. If a mother started 'Hush Little Baby' on G one day, she was very likely to use the same starting pitch on the next day.

There are various complications involved in measuring sung performances. One important issue concerns whether participants used a kind of 'muscle memory' based on past experience singing along to pop tunes in Levitin's study. However, other studies showed that good AP memory can be found in tasks that do not involve singing. For instance, Schellenberg and Trehub (2003) played theme songs from popular television programs from the 1990s (including *Friends*, *E.R.*, and *The Simpsons*) using either their original pitches, or after shifting all the pitches up or down by only one or two semitones. Most people were able to identify the original versions despite the altered versions being very close in pitch, leading the authors to conclude that 'good pitch memory is widespread' (as the paper is titled). Other studies have shown that people can identify audio recordings of familiar pop songs within 200 milliseconds (Schellenberg, Iverson, & McCinnon, 1999)!

In summary, AP may constitute an astounding anomaly of musical ability, or a rudimentary feature of memory, depending on whether you define AP as a labeling ability (rare) or as a feature of music that we can recognize and recall (common). Regardless of all this, nobody disputes that the real meaning of melodies lies not so much with absolute pitches but with their relationships to each other. Next, we turn to ways in which these relationships may be characterized and better understood.

Relative pitch and pitch intervals

Earlier, we mentioned octave equivalence as a way to preserve consistency of musical pitch even when pitch height changes. Octave equivalence is not the only cue that two melodies differing in pitch may be, in fact, the same. Consider that most individuals, even from an early age (i.e., infancy, as discussed in chapter 8), perceive a melody to be the same if someone sings every pitch three semitones higher than the last time the melody was sung. When every pitch is shifted by the same amount, it is called a melodic ***transposition***. For example, if 'Over the Rainbow' is transposed from the key of C to D major (we will describe these concepts in more detail later), the opening phrase changes from C-C′-B-G-A-B-C′ to D-D′-C#-A-B-C#-D′, which changes every pitch in the sequence. Yet most people have no trouble recognizing these melodies as the 'same.' In transposed melodies, even though specific pitches may be changed, the ***intervals*** that pitches form (i.e., relationships between pitches) are retained, and it is these relationships that people primarily use when recognizing a melody.

In fact, these pitch intervals form a set of categories for pitch that is independent of pitch chroma. These relative pitch categories are defined by the intervallic relationships between adjacent (neighboring) pitches or between pitches and a larger context (e.g., a musical key).

In musical practice, one learns of scale step intervals (e.g., major third, perfect fifth). We avoid this terminology (as is common in psychology of music) because these categories assume a specific key. Instead, we refer to semitones (defined earlier), which is a more neutral term. Note that relative pitch categories based on intervals constitute a more abstract way to conceptualize music than absolute pitch. Whereas absolute pitch refers specifically to the 'surface' of the music, which comprises what we perceive directly from the stimulus, relative pitch involves a kind of perceptual abstraction based on comparing two or more pitches (see chapter 7 for a detailed discussion of 'surface' versus 'deep' structure). For the first phrase of 'Over the Rainbow,' the pattern of changes in semitones would be coded as [+12, –1, –4, +2, +2, +1], based on the initial octave jump (C to C′, 12 semitones), followed by a 1-semitone descent (C′ down to B), then a 4-semitone descent (B to G), and so on. Note that the plus and minus signs reflect direction of change, whereas numbers reflect the magnitude of change in pitch in semitones.

One of the fascinating mysteries of music perception is why this abstract way of perceiving music – relative pitch – is so salient to listeners, possibly more salient and important than absolute pitch. It is even possible that absolute pitch perception (defined the traditional way as associations of labels with pitch) may be a source of interference. For instance, musicians without AP labeling ability outperformed those with AP when asked to judge whether two transposed melodies preserved intervallic relationships or not (Miyazaki, 2004).

For most listeners, it is a trivial task to recognize a melody that is transposed. As we will discuss in chapter 8, even infants can do it (Chang & Trehub, 1977). One explanation for the salience of relative pitch may be that melodic intervals imply an underlying harmony, and as such we can use the implied harmonic structure as a way of grouping notes together. Typically, ***harmony*** refers to when multiple pitches are produced simultaneously, as in a barbershop quartet performance. However, a melody can imply a harmony, because listeners associate pitches with each other. The perception of relative pitch may be based in part on the perception of such implied harmonies (Cuddy, Cohen, & Mewhort, 1981), especially for musically trained listeners (Schubert & Stevens, 2006).

Although relative pitch is important, and is perhaps the dominant way in which listeners encode musical pitch, there are limitations of relative pitch perception. Although it is easy to recognize transposed melodies, musically untrained listeners find it almost impossible to identify an isolated melodic interval, even with a great deal of training (Burns & Ward, 1978). The salience of relative pitch may instead be based on the use of contextual information that goes beyond single intervals. The ability to categorize an isolated interval improves considerably when listeners are told to imagine that interval in the context of a larger song. Thus, it is hard to categorize an interval of +3 semitones (a 'minor third') on its own, but easy to imagine it as being the first two notes of 'What Child Is This?' (Smith, Kemler Nelson, Grohskopf, & Appleton, 1994). This tendency probably reflects how we store melodies in our memory when in the context of a familiar song, more so than an alphabet of abstracted pitch intervals.

A common theme that emerges from our survey of absolute and relative pitch is that context matters. The ability to identify absolute pitch of isolated notes is rare, but the ability to recognize whether a song is performed using its original pitches is common. Similarly, the ability to categorize a single pitch interval is rare, but not when the interval is conceptualized as part of a larger whole. Thus, the next sections discuss constructs for understanding the global organization of melodies – that is, the 'context.'

Melodic contour: The importance of direction

Melodic contour refers to the shape of a melody line, in which successive pitches are rising, falling, or unchanging in pitch. It is a reduced form of the code we used before for melodic intervals. For instance, the contour pattern for the first phrase of 'Over the Rainbow' would be coded using the plus and minus symbols from before [+ – – + + +].

There are two reasons why this very simple code is a useful way to characterize the structure of a melody. First, changes in pitch direction are highly salient to listeners – more so than the degree of change in semitones (e.g., Jones, 1987). A performance error that results in a change to the melody's contour is far more noticeable (and far less common) than an error that does not cause such a change. GL, the individual discussed at the start of the chapter, was unusual in that he had difficulty discriminating between short melodies even if the alteration modified the direction of the pitch change, which in turn changes the contour (Peretz, 1993). However, as we will see in chapter 8, even infants can discriminate between two melodies that differ only in one pitch if it alters the melodic contour (e.g., Trehub, Thorpe, & Morrongiello, 1985).

Second, when you hear a melody for the first time and try to remember it, contour is one of the primary features you are likely to retain. Dowling (1978) reported a study in which people heard pairs of melodies and had to determine whether the second melody constituted a transposed version of the first melody or was a different melody. Listeners had a hard time ruling out altered melodies that were 'tonal answers,' which were different melodies (i.e., made up of a different pattern of intervals), but preserved contour and also preserved the overall pitch context known as key (which we describe in the next section). For instance, if one plays [C E D F G] followed by [F A G B C], the two excerpts sound quite similar by virtue of their common contour [+ – + +] and their use of a common musical scale (or key), which here is the key of C, even though the last two intervals differ. Subsequent research has shown that the perception of similarity across melodies may be based on contour patterns across both adjacent and non-adjacent pitches (Quinn, 1999; Schmuckler, 1999, 2010).

Something about the previous paragraph may seem odd. The first five notes of 'Mary Had a Little Lamb' form the same contour [– – + +] as the melody that accompanies the lyric

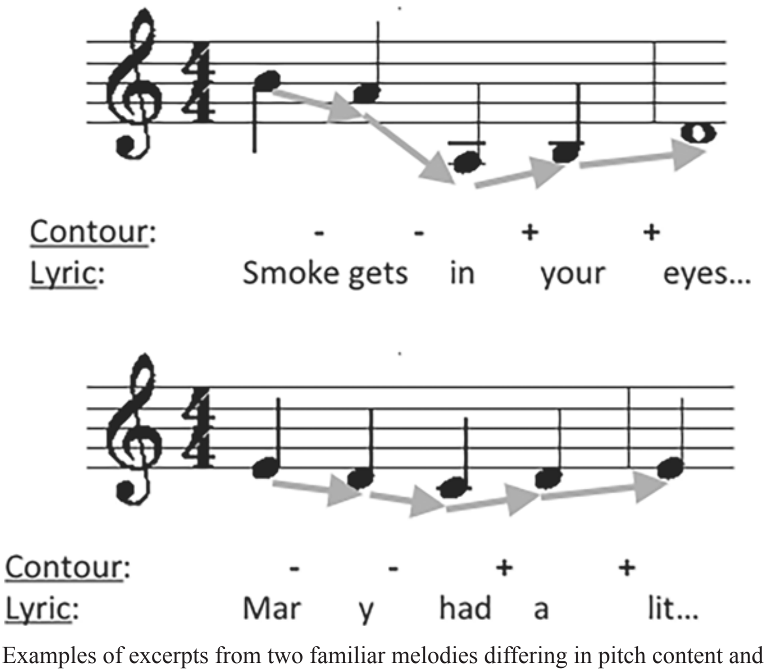


Figure 5.3 Examples of excerpts from two familiar melodies differing in pitch content and lyrics, but sharing the same melodic contour (symbolized by + and – signs).

'Smoke gets in your eyes' from the song of the same name (made famous in the movie *Casablanca*), as illustrated in Figure 5.3. Yet you would not confuse these excerpts, even if they were each played as a solo melody on a piano. That is because the results summarized in the previous paragraph apply primarily to melodies we just heard for the first time (Dowling & Bartlett, 1981; Dowling, Kwak, & Andrews, 1995). As we become increasingly familiar with melodies, our memory starts to store interval information as well. As a result, people get better at ruling out different kinds of alterations.

Thus far, we have made passing references to the musical concept of key. Some readers will be familiar with this construct based on their musical training, others will not. In the next section, we briefly introduce what we mean by key, but the more important focus is on the psychological implications of key for the perception of musical pitch.

Key and tonality: A global context for melody

When you are listening to a melody, it is often easy to hear when one of the notes is wrong, even if you haven't heard the melody before. None of the concepts we have discussed so far accounts for this ability. It comes from the fact that most melodies have a characteristic known as tonality, which means that the pitches in the melody imply a specific musical key.

Scales, key, and tonality

The ***key*** of a melody refers to the specific musical scale on which the melody is based. Melodies that have a perceivable key are said to possess ***tonality***. We focus here on the most common type of musical scale used to establish tonality in Western music genres such as classical, rock, country, and so on, which is the ***diatonic*** scale. Whereas there are 12 possible chromas in Western music, as discussed earlier, in a diatonic scale only 7 of the 12 belong. The simplest example is the C-major scale (major refers to the 'mode' of a key, which is further discussed in chapter 14). If you play a piano (see Figure 5.2) you can set up a C-major scale by using only the white keys. Note that when you do this, the spacing between adjacent pitches is not uniform. If you move from C to D, the separation is 2 semitones (because of the intervening black key), but when you move from E to F, the separation is 1 semitone. This ***asymmetric*** spacing of pitch intervals is an important property of the diatonic scale.

In addition to these first two features of diatonic scales (7 chromas, asymmetric spacing), there is another one that arguably plays the most important role. Different pitches have different roles when a melody implies a diatonic scale, with certain pitches being heard as more important or dominant over other pitches. Thus, a diatonic scale is ***hierarchical***. One division is already implied by how we select pitch chromas. Namely, when a melody is diatonic (e.g., it is in C major), pitches in the scale (white keys) sound like they belong, and the rest (black keys) do not. Beyond that, there is an important distinction between the ***tonic*** pitch of a scale and the remaining pitches. The tonic pitch is a chroma that gives the scale its name. C major takes C as the tonic, D major takes D, and so on. The presence of a tonic pitch is what makes scales diatonic. Perceptually, the tonic is heard as being a kind of central pitch: Melodies that end on the tonic sound finished (as is nicely illustrated in the lyric 'and that brings us back to *do*' in the 'Do-Re-Mi' song), whereas melodies that end on some other note leave the listener feeling uncertain about whether it is really complete.

These issues are central to music theory, but key is not simply a theoretical construct. The effects of tonality are readily apparent to a listener. Beyond the aforementioned fact that wrong (non-tonal) pitches are easy to detect, there is also a salient difference between music that is tonal and music that is atonal. ***Atonal music*** freely uses all 12 chromas, and so does not imply any key. All pitches are equidistant, and there is no hierarchical relationship among pitches. This sort of music was championed by many composers in the early 20th century (e.g., Schoenberg, Webern) as a way of freeing music from the kind of rules it had used previously. Interestingly (and somewhat unfortunately for these composers), atonal music was never widely embraced by the masses. Most listeners prefer ***tonal music*** (i.e., music rooted in a key).

This preference exists even for listeners who have no idea what these terms mean. As hinted at the outset, knowledge of tonality does not require formal training in music theory. The ability to hear whether a pitch fits into the tonal context of a melody comes about even if the listener is not conscious of the melody's key. This is because tonality is learned ***implicitly*** – that is, in a non-conscious way by virtue of exposure to the prevailing music of one's surroundings (Tillmann, Bharucha, & Bigand, 2000). This learning-by-exposure causes us to store knowledge about key as a cognitive ***schema***. Schemas are general frameworks or outlines that we use to structure information in memory (Reber, 1985). Although schemas exist in memory, they are not memories for specific things. Instead, they are memories for how information fits together. For instance, when you reach an intersection while driving in a city, you may assume a stop sign is there even if you don't see one and haven't been to that intersection before, based on schematic knowledge about urban design.

The influence of tonality on pitch perception

The most influential demonstrations of tonal schemata emerged from a paradigm by Carol Krumhansl and Roger Shepard that is highly reminiscent of an anecdote (probably apocryphal, but a good illustration) about the composer Felix Mendelssohn. As the story goes, a visitor was informed that Herr Mendelssohn was in the bath and could not see him. The guest, feeling rather put off, went to the piano and repeatedly played the first seven notes of the scale of C major (up to, but not including, *do*). After several repetitions, Mendelssohn suddenly appeared in his bathrobe – and completed the last note of the scale! The leading tone (or seventh note of the diatonic scale) sets up such a strong expectation to resolve to the tonic that Mendelssohn could not bear to leave the scale incomplete.

The scientific variation on this theme is called the ***probe-tone technique***. First, one hears a sequence of notes or chords that sets a context. This can be as simple as an incomplete scale (as in the Mendelssohn anecdote). Next, one hears a 'probe' tone, which can be any one of the 12 chromas. One rates how well the probe tone fits or completes the context. The first study to report this technique revealed that listeners could hear hierarchical relationships among notes that are based on the diatonic scale implied by the context (Krumhansl & Shepard, 1979). Subsequent research showed that this effect generalizes even to young children (Cuddy & Badertscher, 1987; for a review, see Krumhansl & Cuddy, 2010).

But tonal schemas do not just influence how we perceive single tones; they also influence how we perceive relationships between tones. This result came from a follow-up study in which participants heard two probe tones, and indicated how similar they sounded to each other (Krumhansl, 1979). In such a task, there are hundreds of similarity ratings to analyze, and it is usually impossible to plot these relationships on a standard scatterplot or as means across treatment conditions. A common way to understand the relationships that people perceive in such complex contexts is to use a technique known as ***multidimensional scaling***. It involves a computer algorithm that generates a geometric shape that best represents the relationships across all pairs. The basic assumption of this algorithm is that perceptual

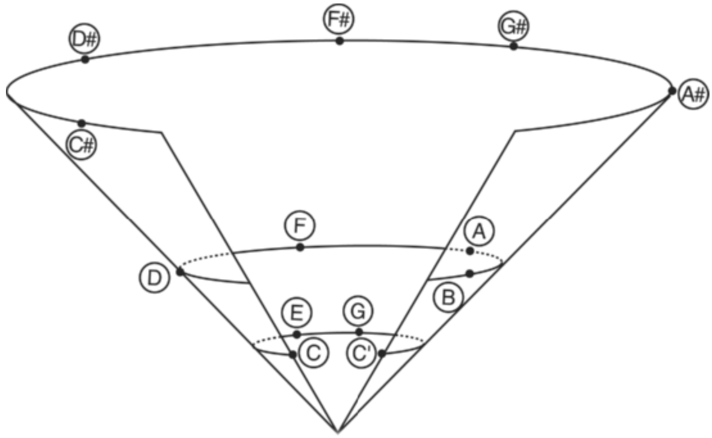


Figure 5.4 Conical representation of similarity ratings for pitches within a C-major context.

Source: Krumhansl (1979, Figure 3, p. 357), reprinted with permission from the author and Elsevier.

experiences that are similar should be represented as nearby points in space. Although a multitude of shapes could be generated, the computer algorithm attempts to provide the simplest possible image for the observer to understand. One simple constraint is to find the representation with the fewest spatial dimensions, with three or fewer being ideal.

The similarity ratings that Krumhansl collected best conformed to a conical shape, as shown in Figure 5.4 (described by the author as a 'slightly idealized' scaling solution). Similarities are shown as geometric distances. Points closer around the surface of this shape represent pitches that sound similar within a tonal context. This particular figure shows where pitches lie when the context is of a C-major scale, though in principle this shape could apply to any major key (you would simply change the pitch labels associated with different points). Thus, when a C-major context is followed by the pitch pair [C G], a participant would be likely to rate the pair as highly similar, whereas the pair [C D#] would be rated as highly dissimilar. This reflects one characteristic of the cone: Pitches within the scale cluster together (toward the bottom, reflecting a kind of gravitational pull), with pitches outside the scale at the top layer.

Note that the cone also narrows as you move from top to bottom. As such, the lowest level (the C-major triad) forms a semicircle with a small diameter. This means that all pitches within the established key were heard as similar to each other. By contrast, the top level forms a semicircle with a wider diameter. This means that pitches outside the tonal context were heard as being dissimilar not only to in-key pitches but also to each other. Thus, pitches outside the scale sound as if they are disconnected and free-floating, rather than forming a distinct coherent group. This may underlie the experience some listeners have of atonal music as being hard to organize perceptually.

Name that tune: Memory for melodies

A lot of the research we have discussed so far has involved melodies created by experimenters to manipulate certain features of melodies (e.g., melodic contour) while holding other features constant (e.g., tonality). Although it makes sense to do this from the perspective of the scientific method, there are limitations to such artificial melodies. A particularly salient limitation is that these melodies tend not to be very 'catchy.' As a result, they may not provide a good basis for understanding long-term memory for melodies. Listener recognition for such artificial melodies is surprisingly poor, and much poorer than our ability to recognize familiar melodies (Halpern & Bartlett, 2010). Thus, we here consider the factors that contribute to our ability to recognize familiar melodies.

One critical factor is the role of ***musical phrases***. Musical phrases are sequences of notes within a melody that are heard as belonging to the same musical group. Notes within a phrase are perceived as coming to some kind of closure by the end of the phrase. The degree of closure can vary (see chapter 7 for more discussion on the basis of closure). Phrases are often categorized into those that function like questions, for which the ending seems to require further completion by a subsequent phrase (think of the first line: 'Twinkle, twinkle, little star'), and those that answer, or come to a resting point ('How I wonder what you are'). Studies that have explored how long it takes listeners to identify a tune have found that a decisive point in recognition comes at the end of the first phrase; at this point, your ability to 'name that tune' improves dramatically (Dalla Bella, Peretz, & Aronoff, 2003; Schulkind, Posner, & Rubin, 2003). In addition, certain characteristics of pitch within phrases can be particularly illuminating, such as repetitions of patterns (Schulkind et al., 2003).

But the cues for a melody's identity do not just exist within the melody itself. Memory often involves some kind of mental comparison. When remembering where you parked today, you must be able to filter out competing memories for the parking spaces you used over the past week. In the same way, musical features that facilitate memory can also be found when comparing the structure of a given melody to other melodies one might know. Müllensiefen and Halpern (2014) demonstrated this by analyzing features of melodies in comparison to a larger corpus of melodies, which included popular songs. For instance, the authors found that melodies with large interval jumps (they use 'Under the Boardwalk' by the Drifters as an example) facilitated the feeling of recognition, even if a listener hadn't heard it before. Large jumps like those used in this tune constituted a distinctive feature of the melodies in the corpus because most melodic intervals are small.

Involuntary memory for melodies: Earworms and priming

Although it is natural to consider memory for melodies to be effortful (as in the classic US television show *Name That Tune*), there are many cases where retrieval occurs automatically, perhaps even when we do not want it! This kind of experience has been referred to as having an ***earworm*** – a form of involuntary musical imagery colloquially known as when a tune gets 'stuck in your head' (this term comes from German, and is based on a rather grotesque analogy between the tune stuck in one's head and a worm burrowing into one's ear).

Based on a recent estimate, 72% of us may experience multiple earworms each week, and many of us even experience them on a daily basis (Halpern & Bartlett, 2011). The experience of having an earworm is not necessarily unpleasant: In fact, about 64% of earworms are reported as either neutral or pleasant. In a detailed reflection on his own earworm experiences, Brown (2006) refers to earworms as comprising looped fragments from musical pieces, rather than a piece in its entirety, with the boundaries for the beginnings and ends of fragments being 'fuzzy.' Not surprisingly, recent exposure to music is a good predictor of whether that music may become an earworm (Brown, 2006; Williamson et al., 2011). Based on their analysis of two large online surveys (with over 18,500 respondents), Williamson and colleagues found that two ways people respond to earworms is by engaging (e.g., listening to the whole tune to complete the fragment that 'stuck' in their heads) or distraction (such as talking, or thinking of another melody to 'cure' the earworm). Beyond these themes that emerged, however, there were also finer variations in responses to earworms (Williamson, Liikkanen, Jakubowski, & Stewart, 2014).

Memory need not always involve a conscious recollection (explicit memory). Sometimes memory leads to subtle effects on our ability to perceive new information, leading to an effect known as ***priming***. Priming occurs when the perception of a stimulus makes related items from memory more easily accessible. As a result, the perceptual system can more easily encode these related items for a brief period of time. Priming depends on a process psychologists call ***implicit memory*** (cf. our earlier discussion on the implicit role of tonal schemas). Implicit memories influence how we perform tasks, but are not typically accessible to consciousness. In fact, it is often difficult to verbalize an implicit memory. For instance, we know how to tie our shoes based on memory, but most people find it very difficult to describe how this task is performed!

Many parallels have been drawn between priming effects in language and in music (for a review, see Tillmann & Bigand, 2002). In semantic priming in language, the recognition of a word can be enhanced if relevant material is presented earlier. For example, if someone is shown the word 'donkey,' the speed with which they can distinguish 'horse' from 'hoarse' is increased (McNamara, 2005). Similarly, in musical ***harmonic priming***, processing of a chord tends to be faster and more accurate when preceded by a chord that is harmonically related or 'schematically probable,' compared to one that is not (Justus & Bharucha, 2001; Tillmann & Bigand, 2002). This effect has been demonstrated in studies that measured behavioral responses as well as experiments that employed brain scanning with functional magnetic resonance imaging (Tillmann, Janata, & Bharucha, 2003). A variation on this theme, ***repetition priming***, happens when repetition of the identical stimulus enhances your perception of it, as in the case of a melody with repeating pitches (Hutchins & Palmer, 2008).

As we have now learned, even a short sequence of tones can bring about a cascade of mental processes, including pitch processing, use of tonal schemas, priming, and recognition based on long-term memory. Yet the examples we have discussed so far lack an important connection to the music we usually encounter. So far, we have discussed melodies as if they exist in isolation. Although melodies are sometimes produced this way, more often we listen to music comprising many instruments and voices. This combination of sound sources introduces a new challenge to the perceptual system that we consider next.

Hearing multiple melodies

Auditory scene analysis for music

The music we listen to rarely features an isolated melody line. Typically one hears a singer accompanied by several instruments, or an instrumental piece in which one instrument carries the melody and the rest accompany it. Finally, there is music featuring multiple melody lines that interact, such as in Bach's polyphonic works or in the Beatles' 'She Loves You,' in which the parts sung by Paul McCartney and John Lennon function like two intertwining melodies. In order for this to happen, our brains have to carry out two complementary functions. On the one hand, we need to ***segregate*** sounds that belong to different parts (e.g., the violin as opposed to the trumpet). On the other hand, we need to ***integrate*** sounds that belong to the same part (e.g., keeping together violin tones that may vary in pitch, loudness, duration, and subtle aspects of timbre). Importantly, all this needs to happen *before* we can make sense of a melody, as we will show later.

The ability to perceive the auditory signal as coming from a set of distinct sources was dubbed ***auditory scene analysis*** in a classic book of the same name by Albert Bregman (1990).

Auditory scene analysis is assumed to occur before we can recognize melodies, infer tonality, and so on. As such, complex schematic representations for tonality may not have a strong influence on whether you can perceptually separate the two melodic lines of 'She Loves You.' Although auditory scene analysis arises from basic acoustic features, it is an error to assume that it is therefore 'easy' for the auditory system to accomplish. Certainly, our experience of analyzing sounds in this way feels effortless. But the computational demands for the auditory system are significant. As Bregman points out, the signal reaching our ears combines all of the sounds in our environment. In other words, there is no segregation in the signal. Bregman compares the task to a visual analogy: Imagine trying to determine how many motorboats are in a lake, using only a pattern of rippling water on the side of the lake (pp. 5–6).

Bregman adapted the concept of ***Gestalt grouping principles*** to explain how auditory scene analysis works. These principles characterize the circumstances under which we perceive small-scale perceptual features into organized wholes (or objects). In vision, these features may be lines, surfaces, or shapes. In audition, features may be events such as musical notes and chords.

The word *Gestalt* is German for 'whole form.' In other words, gestalts are the organized structures that emerge from the seemingly disparate physical stimuli in our environment. This school of thought was initiated by Max Wertheimer (1880–1943) and extended by others, notably Kurt Koffka (1896–1941) and Wolfgang Köhler (1887–1967). Gestaltists believed that perceptual systems (vision and audition) prioritize the experience of wholes (or patterns) rather than the parts that make them up. In other words, we hear melodies, not pitches. The original Gestaltists focused primarily on the role of grouping principles for vision, but Bregman expanded this view to audition. With respect to grouping principles in the auditory domain, one principle stands out from the others with respect to its influence over perceptual organization: the *principle of proximity*.

The power of proximity

In vision, the ***principle of proximity*** states that (all else being equal) the elements that are near to one another in space tend to be perceived as a group. For audition, one is tempted to translate visual space in a simple way to auditory space. Unfortunately, the auditory system does not localize sounds precisely in space, in comparison to our ability to locate objects using vision. As such, the kind of proximity that seems to dominate auditory perception has to do with closeness in pitch height. Pitches that are close in proximity tend to be heard as if they belong to the same source. This makes sense, given that there are physical constraints on the range of pitches one can produce with one's voice or on an instrument.

W. Jay Dowling demonstrated just how important pitch proximity is to melody perception and recognition (Dowling, 1973). He investigated the conditions under which listeners are able to separate two ***interleaved melodies***. Interleaving was accomplished by playing alternate tones between two familiar melodies (i.e., first tone of melody A, first tone of melody B, second tone of melody A, second tone of melody B, and so on). Here is a visual analogy of interleaving, employing interwoven sequences of letters:

INTERLEAVED

TWOMELODIES

When interleaved, the result would be:

I T N W T O E M R E L L E O A D V I E E D S

As you can see, it is incredibly hard to make out the original words!

A similar effect happens in music. Dowling showed that when two interleaved melodies share the same pitch range, it can become *impossible to identify* even an extremely well known melody like 'Mary Had a Little Lamb.' The reason is that all the pitches are heard as belonging to the same source. As a result, listeners cannot perceptually segregate the melody in order to recognize its contents. However, when the interleaved melodies are played in different pitch ranges, the two groups of tones are heard as two distinct melody lines and recognition can occur.

Whereas Dowling's study shows the importance of pitch segregation in recognizing a melody, a related study by Diana Deutsch shows how the integration of pitches within a melody is also dependent on their proximity. She took highly familiar melodies and created ***octave-scrambled melodies*** (these can be heard in Deutsch's 1995 CD *Musical Illusions and Paradoxes*, Track 19–23, 'mysterious melody'). Octave-scrambled melodies preserve the notes and rhythms of a familiar melody, but play each note in a different octave (most effective if using the full range of the keyboard and selecting octaves without any pattern). The tones do not cohere into a melody, but seem to 'fly apart'! As in the Dowling study, this manipulation of pitch height prevents recognition. Only now it is the pitches *within* the melody, rather than distracting *incidental* pitches, that are made to sound like they belong to different sources.

Note that these examples are not just the tricks of experimental psychologists. Similar effects can be found in musical works, in which alternating between high and low registers may create the impression of two melodic lines played by a single instrument (virtual polyphony). One famous example is Bach's solo violin works (e.g., Davis, 2006). In these works, rapid shifts in pitch height are used to give the sense that more than one violin melody is occurring. A related example exists in expert yodeling, in which the modal and falsetto notes of a male singer appear to form two distinct melody lines.

Other grouping principles

Although pitch proximity seems to play a dominant role in grouping, there are other Gestalt principles that appear to play analogous roles for vision and audition. One is the ***principle of similarity***. Similarity relates to whether perceived elements belong to the same category. This principle therefore relies more on subjective criteria than does proximity. In the domain of vision, similarity is often based on shape or color. In music, the standard analogue is timbre. We often organize the input of individual sounds into organized groups such as 'melody' and 'accompaniment,' 'strings,' 'winds,' 'brass,' and so on. In such circumstances, it is often timbre that serves as the quality that allows the listener to segment sounds into different instrument groups.

In some cases, timbre-based similarity may supersede pitch proximity. A thought-provoking auditory illusion discovered by Diana Deutsch, called the 'scale illusion' (1975), makes this point while also demonstrating the power of pitch proximity. In Deutsch's study, participants listened to a musical sequence through headphones. The sequence was based on the C-major scale, with the notes of the ascending and descending scale alternated simultaneously between the left and right ears, producing two rather jagged melodic lines played to each ear (as illustrated schematically in the top part of Figure 5.5).

Participants were asked to give a verbal report of what they heard through the headphones, and to sing what they heard in each ear. Not one of the 70 participants reported hearing the musical pattern that was actually played! As illustrated in Figure 5.5, most participants perceived smooth ascending and descending contours of the scale, and associated the higher

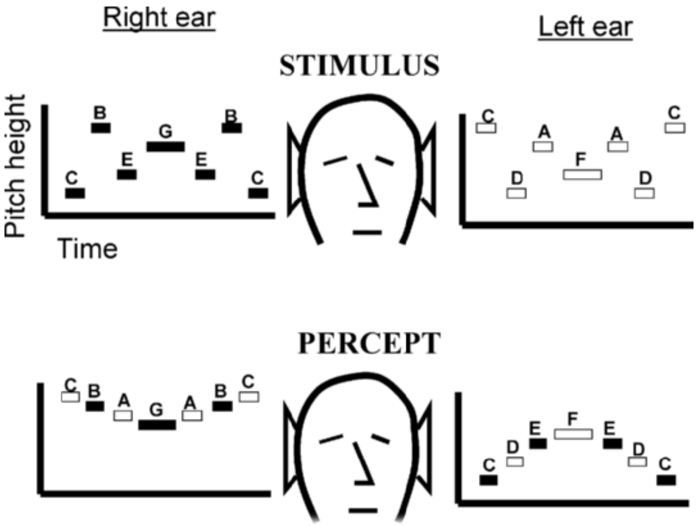


Figure 5.5 An illustration of the scale illusion (Deutsch, 2013). Rectangles denote tones varying in pitch and timing. Filled rectangles denote tones presented to the right ear (through headphones), and open rectangles denote tones presented to the left ear.

tones with the right ear. The most common percept was hearing half of a descending C-major scale, followed by an ascent back to C in the right ear, with the opposite percept (half an ascending scale followed by a descent to C) in the left ear. In other words, the clearly discriminable differences in spatial location were overcome by listeners' tendency to hear patterns based on smooth melodic contours determined by more proximal pitches.

So far, the scale illusion suggests that grouping based on pitch prevails over grouping based on spatial location. However, this is not always the case, as other features can be altered to prevent the scale illusion from taking place. In particular, if the sounds presented to different ears are also associated with sufficiently different instrumental timbres, the participant is more likely to hear locations of the sounds as they were originally presented. In this case, similarity of timbre, when combined with spatial location, overcomes pitch proximity. More broadly, if the tones are played in sufficiently different pitch ranges, sufficiently different timbres, or if the tone sequence does not imply an orderly scale, the effect is likely to be weakened or absent, as the perceptual principles do not act alone, but often act cooperatively or in competition with one another.[2]

Another well-known Gestalt principle is the ***principle of closure***, the tendency for our perceptual system to fill in missing parts of an easily recognizable aggregate (e.g., when the corner is removed from a square, you still perceive the square). Experimental evidence for closure in music was reported by DeWitt and Samuel (1990). The authors tested whether an effect similar to phonemic restoration in language occurs in music. Phonemic restoration (Warren, 1970) is a powerful effect in which listeners continue to hear intact speech even when a phoneme is replaced by noise. Rather than hearing the noise where it occurs, the listener experiences the speech sound as if continuing *through* the noise. This effect is analogous to 'filling in a gap' visually, and tends to be stronger at the end of sentences than at the beginning, suggesting that the process of filling in requires knowledge of the context, which is referred to as 'top-down processing.' DeWitt and Samuel (1990) explored whether a general ***perceptual restoration*** occurs in music, wherein listeners replace missing sounds based on their expectations. In one experiment, DeWitt and Samuel created 10-tone scales (octave scale plus next two tones) with one tone either present but accompanied by noise, or replaced by noise. Listeners had greater difficulty discriminating whether the tone was present or absent if the distorted tone was the expected pitch that would complete the scale (as opposed to a random pitch). They were also more likely to 'restore' a missing pitch as the number of tones of the scale before the distorted tone increased. Thus the findings showed 'an increase in restoration associated with increased musical expectations' (p. 141).

A general theme of this section is that our ability to separate the sounds of different instruments while listening is based on general-purpose perceptual principles that extend to other kinds of auditory perception (e.g., speech) and are even analogous to vision. This contrasts with the more music-specific structure of tonal schemas discussed earlier. This makes sense because in order to recognize a melody, one must *first* be able to separate the melody line so as not to base tonal judgment on a confusing mixture of different instrumental parts. We can therefore see how the different processes discussed in this chapter flow in a general sequence, with auditory scene analysis (when needed) coming first.

Neural bases of pitch and melody perception

In this final section, we review selected examples from the broad and growing neuroscientific literature on musical pitch. In the interest of brevity, we focus here on the most music-specific processes discussed so far: The perception and memory of musical pitch, key, and melodic contour. Figure 4.2 may be useful in locating brain regions referred to in the following discussion.

Chroma

A fundamental assumption about musical pitch is that it is multidimensional, with the two most basic dimensions being pitch height and chroma (as discussed before). Although this distinction clearly relates to the notation categorization system that we use (which separates octaves from chromas), it is less clear whether neural processing of auditory pitch would reflect such a separation.

An ambitious study by Warren and colleagues sought to test this possibility (Warren, Uppenkamp, Patterson, & Griffiths, 2003). They presented participants with pitch patterns in which the variability across pitches could be based on changes to both height and chroma, or could be based on just one dimension. In doing so, they had to face a daunting challenge. As we know from the pitch helix, it is easy to vary height while keeping chroma constant – you just switch octaves. By contrast, it is not clear that one can change chroma while keeping height constant. What the authors did was vary the amount of energy across different frequencies of complex tones (cf. chapter 2), so that the average frequency across all the harmonics remained constant. This leads to an ambiguous sense of what octave a given chroma is in, thus reducing the salience of changes in pitch height across successive tones. Through these manipulations the authors identified distinct brain regions that responded to variability in each dimension. Interestingly, separation of height and chroma was not found in the *primary auditory cortex*, but in regions in the *belt* surrounding it (see chapter 3). Whereas changes to chroma led to activations *anterior* to (i.e., in front of) the primary auditory cortex, changes to height led to *posterior* activations.

One issue that has intrigued researchers in music cognition is the degree to which the brain possesses fixed chroma-specific pitch categories based on absolute pitch. The dominant view, as discussed earlier, is that most people do not remember absolute pitch (i.e., they do not recognize chromas independent of their context). From a neurophysiological perspective, the rarity of absolute pitch presents a puzzle. After all, the cochlea of the ear and the auditory cortex share what we call a ***tonotopic*** representation of pitch (see chapter 3), which is chroma-specific, similar to a piano. By contrast, we do not have evidence for brain areas that respond to specific musical intervals (relative pitch). Why, then, is absolute pitch (defined as a specialized labeling ability) so rare?

One possible explanation is that pitch is processed differently in the brains of individuals with absolute pitch compared to individuals who lack this ability. With respect to anatomy, musicians with absolute pitch have an asymmetry between the left and right hemispheres that prioritizes the left hemisphere. This asymmetry is found in an auditory association area known as the ***planum temporale***, positioned towards the rear (posterior) of the temporal lobe (Schlaug, Jäncke, Huang, & Steinmetz, 1995). This finding is consistent with other research suggesting that the ***lateralization*** (i.e., the degree to which a neural function is exclusive to the left or right hemisphere) of musical pitch varies with training. For instance, whereas for most people musical pitch is processed predominantly in the right hemisphere, with the left dominant for language, trained musicians may also show left dominance for music (Bever & Chiarello, 1974). Thus the left hemisphere may be dominant for auditory processing that involves the application of internalized categories, such as absolute pitch. Yet another, more recent proposal is that the right hemisphere is optimized for analyzing the spectrum of a sound (which yields its pitch; see chapter 2), whereas the left hemisphere is optimized for perceiving rapid temporal fluctuations (common in speech; Zatorre, Belin, & Penhune, 2002).

A more recent study using fMRI compared neural responses in the brains of absolute pitch possessors to musicians who lack this ability (Loui, Zamm, & Schlaug, 2012). The authors found greater activation in the auditory regions of the brain for absolute pitch possessors, and ran further statistical analyses of ***functional connectivity*** (whether activity in one brain region predicts activity in other regions) that suggested greater interconnectivity across auditory brain areas for those with absolute pitch. A surprising aspect of this study is that the task the authors used had nothing to do with pitch categorization, but involved rating the degree of emotional arousal in musical excerpts. Thus, the brains of absolute pitch possessors may process musical pitch in a different way from other brains, regardless of the task.

Key

One of the most striking characteristics of musical pitch processing is the tendency to categorize pitch with respect to the surrounding tonal context (Krumhansl, 1990). In fact, there seem to be specific regions that serve the recognition of the sequences of notes that make up a scale. A study using fMRI attempted to identify brain regions responsible for detecting when a single tone ‘pops out’ based on its divergence from a tonal context (Janata et al., 2002). The researchers simply inserted tones that did not belong to the key, such as an F# in the key of C major. The study found consistent activation in the *superior temporal gyrus* (see Figure 4.2), an area of general importance for pitch perception (Peretz & Zatorre, 2005), more so for the right than left hemisphere. However, the main focus of the authors was on what part of the brain ‘tracks’ changes in tonality, which they emphasized as being the *rostromedial prefrontal cortex*, an area just behind the center of one’s forehead.

Later research using fMRI showed that this area’s correlation with tonal motion is greater for music that evokes strong autobiographical memories (Janata, 2009). So when a familiar song suddenly sends your mind back to the first week of college, your brain is responding more strongly to its tonality than when listening to a similar song that has no such strong association for you.

Other research concerning the way in which tonality makes certain pitches 'stand out' from the rest has employed the use of the ERP technique. This research has revealed changes in electrical activity when a pitch does not match the overarching scale context. Interestingly, the brain responds to out-of-tune notes (which do not require a sense of scale, but only require an understanding of basic pitch categories) more quickly than unexpected notes that are appropriately tuned (Brattico, Tervaniemi, Näätänen, & Peretz, 2006). This finding suggests that the interpretation of pitch within a musical scale may constitute a higher-level cognitive task relative to detecting a mistuned note. A more recent study using fMRI revealed that out-of-tune judgments may be influenced by the context surrounding the mistuned note (Royal et al., 2016). Generally speaking, pitch judgments involve a broad network of brain areas, not surprisingly dominated by bilateral activations in the superior temporal lobe. However, when you compare two pitches within the context of a melody, as opposed to comparing isolated pitches, there is an increased amount of activity in the right inferior parietal lobule, an area of the brain that has traditionally been associated with comprehending the location of objects in space.

It is possible that particular neural processes are reserved for perceiving and/or remembering a pitch sequence as an integrated whole and for registering deviations from this. Much of the case study research reported by Isabelle Peretz and colleagues addresses this issue. In the paradigms used by this group, participants are typically presented with melodic sequences in pairs, in which the second sequence is identical to the first with the exception of one note, which may be altered in pitch or duration. Findings from this research suggest that damage to the right temporal lobe selectively disrupts the ability to detect deviations in melodic contour (Liégois-Chauvel, Peretz, Babaï, Laguitton, & Chauvel, 1998; Peretz, 1990). Similarly, imaging techniques suggest a separation in the regions where melodies are recognized by overall pitch contour and where they are recognized by sequences of specific interval relations alone. The former tends to be dependent on a right-hemisphere activity, while the latter seems to involve regions in both hemispheres (Peretz & Zatorre, 2005). Peretz (1990) interpreted this finding as supportive of the view that the right hemisphere is involved in more 'holistic' processing, whereas the left hemisphere is specialized for finer details.

Integration

As we opened the chapter with a case study, we conclude our discussion in a similar fashion. The case of Rachael Y., described by neurologist Oliver Sacks (2007), serves as a poignant example of the importance of our ability to integrate rich musical works into Gestalt wholes, an ability most of us take for granted. Rachael Y. was an accomplished middle-aged composer and performer when she suffered severe injuries to the head and spine following a serious car accident. (The locations of the injuries were not specified in Sacks' account.) Upon recovering from a coma that lasted several weeks, she found most abilities such as speech to be intact, but noticed a change in her perception of music. She described her experience of the first piece she heard following her recovery from the coma (Beethoven's opus 131):

> When the music arrived, I listened to the first solo phrase of the first violin again and again, not really being able to connect its two parts. When I listened to the rest of the movement, I heard four separate voices, four thin, sharp laser beams, beaming to four

> different directions. Today, almost eight years after the accident, I still hear the four laser beams equally … and when I listen to an orchestra I hear twenty intense laser voices. It is extremely difficult to integrate all these different voices into some entity that makes sense.
>
> (p. 113)

As we have discussed, melody is not perceived in isolation. It does not only require the perception of a horizontal sequence, but the complex interweaving of all the parts including the vertical axis (harmony). Having lost the basic ability to integrate the many rich parts of music into a coherent whole, Rachael Y. experienced music as quite unpleasant and chaotic, requiring 'a great cognitive effort to hold the strands together' (Sacks, 2007, p. 116).

Coda

In many ways, the perception of melody is central to our experience of music. GL's case illustrates the dramatic loss of ability to recognize the storehouse of tunes we accumulate throughout our lives. Without the ability to conceptualize tone sequences as gestalts, and to hear the tones as part of a coherent tonal system, music becomes meaningless. At the same time, you may have noticed that the term 'melody' in this discussion has almost exclusively referred to pitch. It is common for those in music cognition to use the term melody in a way that excludes rhythm. The assumption here is that rhythmic relationships contribute independently to musical experience. But do they? We consider this, and other issues related to rhythm, in chapter 6.

Notes

1 See also Satoh, Takeda, and Kuzuhara (2007).
2 Not all participants heard exactly the same percept. Several examples of how the pattern is perceived are shown in the liner notes to Deutsch's (1995) CD. Some participants hear only a single pitch (same pitch 'heard' in both ears), so that two tones are not even perceived. Despite some variation in responses, none of the reports corresponds to the actual notes played in each ear.

6 Perception of musical time

So you think you can't dance?[1] Though many of us may feel shy about our ability to shine on the dance floor, most people are able to synchronize movements of their body with the dominant pulse, or 'beat' of the music. Yet, as is the case with the ability to perceive musical pitch, the ability to find the beat may be diminished or even absent in some people. A compelling case of this sort was reported by Phillips-Silver and colleagues (2011). Mathieu exhibits 'beat deafness.' Although Mathieu can tap to a regular metronome click, he is unable to find the beat when listening to more complex musical rhythms. For instance, Mathieu does not hear the prominent alternation between strong and weak beats in *meringue* dance music. As such, his movements while dancing are largely independent of the music's temporal structure. Further research suggests that a small segment of the population may be like Mathieu: unable to follow musical rhythms despite having intact musical pitch perception (Launay, Grube, & Stewart, 2014; Sowiński & Dalla Bella, 2013).

Mathieu's problem involves finding structure in musical time. This is a critical problem because the perception of temporal structure allows us to dance, sing along with others, perform rounds, and in many ways enjoy the richly social and emotional significance of music. The importance of time in music goes even beyond all of this.

Consider what music would be like if every pitch of a symphony or song was played at once! This rather extreme example highlights the importance of time, if only as a way of separating tones from each other. But the ordering of each tone in time clearly also matters: A scrambled version of a tune as simple as 'Mary Had a Little Lamb' would be unrecognizable. Moreover, we know that tone durations also matter. Consider the first several tones of 'Mary Had a Little Lamb' and 'The First Noel.'[2] When sung in the same key, each melody has the same succession of pitches; it is only the timing of the pitches that differentiates them. Moreover, there are forms of music that are solely or mostly defined by time patterns. Consider drum circles, rap music, or 'patter' songs (the classic Broadway musical version of rap – think 'Why can't the English teach their children how to speak?' from the musical *My Fair Lady*).

Despite the importance of pitch, music does not require variations in pitch. Drummers can express a range of emotions simply by varying produced temporal patterns (Laukka & Gabrielsson, 2000). In fact, one might be tempted to speculate that the temporal domain is even more important to music's evolutionary history than pitch. Benefits of music for our ancestors may have been based on its ability to facilitate synchronization, whether on the level of groups (Huron, 2003) or in mother–infant relationships (Dissanayake, 2000; see chapter 8 in this volume). In any case, it is important not to discount the role of rhythm and timing in music.

The case of Mathieu shows that we should not take for granted an apparently simple task like tapping to a beat. As with pitch (chapter 5), musical time is multifaceted, and thus involves more complex perceptual processes than we may realize. Our discussion of musical time in this chapter reflects this multifaceted nature. We first discuss rhythmic patterns, which arise from the timing of musical events in music. From rhythmic patterns, the listener derives more abstract characteristics of temporal structure. One of these is the perception of rate, or tempo, which we consider next. Finally, we consider the role of meter, another form of musical time that is associated with perceived regularity. Like tonality, which we discussed in chapter 5, meter is an abstraction that helps us ascertain structure and interpret rhythms. The deficit exhibited by Mathieu likely represents a failure to perceive meter.

Rhythmic patterns

What, exactly, is rhythm? One of us (PQP) attended a workshop on rhythm held by the neuroscientist Aniruddh Patel. As part of the workshop, the participants (all researchers in music cognition) offered their definitions of rhythm. Nearly every definition was strikingly different! The diversity found in the responses highlights an important point. In comparison to a construct like tonality, for which we have a reasonable degree of consensus, the concept of rhythm resists simple definitions. Yet, in the interest of clarity, we offer one here that is used in many (though not all) theoretical discussions of musical time.

We define ***rhythm*** as a temporal pattern created by tone onsets, or the onsets of other musical 'events', which may be notes, chords, dogs barking, and so on. The time between onsets forms what is called an ***inter-onset interval (IOI)***, and rhythmic patterns are based on these. It is important to notes that tone ***durations*** (the time from beginning to end), though important, do not determine rhythms. Take, for instance, the melody shown in Figure 6.1. Though the version in Figure 6.1a includes many durations that differ from the melody shown in Figure 6.1b, both rhythms would sound the same. Duration is important in music because it can add emphasis to certain notes (e.g., the note beginning Figure 6.1a); however, durations are not the primary generator of rhythms.

Having identified the importance of onsets in rhythms, consider how onsets come to form a pattern. To say that a stimulus has a pattern suggests that there is some recognizable structure to it. If you throw several pots down a wooden staircase, you will hear a series of inter-onset intervals, but you probably will not perceive a rhythmic pattern. So what gives a rhythm a perceivable pattern?

How we perceive rhythmic patterns

One important characteristic of rhythms is that they are based on ***relative time*** rather than ***absolute time***. Absolute time is time as it appears on a stopwatch, a time-span with no comparison. For instance, we can measure the first inter-onset interval for the melody in Figure 6.1a as being 600 milliseconds. This time measurement is not influenced in any way by the timing of surrounding inter-onset intervals. By contrast, relative time involves the relationship between a time interval and its surrounding context. Terms from music use relative time. One would call this first note a 'dotted quarter note,' or 'dotted crochet.' This term is based on the relationship of this first tone's timing to the surrounding context.

There is a practical importance to relative time. In music, rhythms are supposed to remain constant even when they are sped up or slowed down. Rhythm thus cannot depend on absolute time because the absolute time of every note changes when the overall rate changes. If you take the rhythm associated with the short musical couplet 'Shave and a hair cut (pause) two bits,' and speed the whole thing up, most listeners can still recognize it. There is an analogy that can be drawn to pitch here. As discussed in chapter 5, melodies are defined

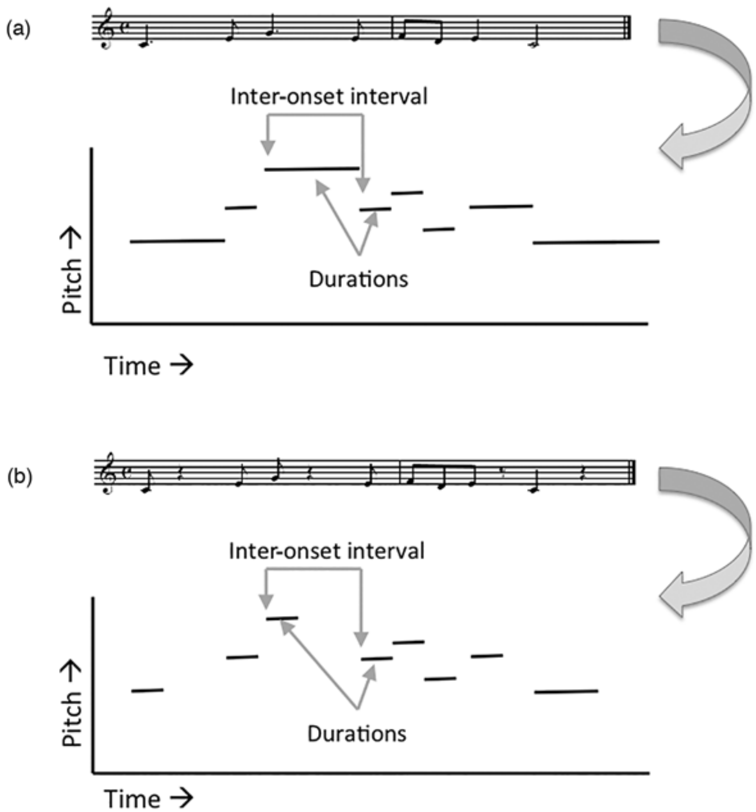


Figure 6.1 Two melodies (a versus b) comprising tones of differing durations that yield the same rhythm based on the timing of onsets. Plots below each melody display pitch and duration values in a graphical format where each dash is a note event. Inter-onset intervals span from the start of one note event to the next; the interval between notes two and three is highlighted. Durations span from the start of one note event to the end of the same event, reflected by the width of each dash.

Source: Copyright © Peter Pfordresher.

based on relationships among pitches. As a result, when we shift all the pitches in a melody up or down by the same amount (i.e., we transpose the key), the melody remains recognizably the same because relationships among pitches in the melody remain unchanged. We do something similar when we speed up or slow down a rhythm.

The fact that rhythms are based on relative time leads to the conceptualization of rhythmic relationships as ratios. The ratio formed by two adjacent time-spans is their ***serial ratio*** (Jones, 1976). The serial ratio metric offers one indication of a rhythm's complexity – that is, how easy or hard it is to follow along with and reproduce that rhythm. Simple rhythms come from simple ratios, which are those that can be reduced to an integer value. Such simple ratios are most likely to elicit the perception of a pattern. For instance, if one time interval is exactly half the length of an adjacent interval, regardless of their order, that ratio can be expressed as a 2:1 ratio. This is the so-called 'swing ratio' used in jazz. However, if one time interval is 500 milliseconds and the adjacent interval is 157, the ratio of 3.18471…:1 is considerably more complex. The simplest ratio is a 1:1 ratio, referred to as an 'isochronous' ratio. The serial ratio offers a prediction concerning what kind of temporal patterns may sound rhythmic. At the same time, as we will see, the serial ratio does not offer a complete account, and thus must be considered an approximation.

The fact that serial ratios can predict rhythmic complexity was demonstrated in a classic study by Povel (1981). Participants tapped rhythmic patterns of alternating inter-onset intervals that varied in their ratio relationship. People did not do this task perfectly, and their errors indicated a tendency to drift toward a simple 2:1 ratio. At the same time, this study illustrated a limitation of the serial ratio construct. When participants were finally asked to tap a 2:1 ratio, none of them could precisely do this! Instead, participants drifted to a ratio slightly under 2:1, in the direction of the simpler 1:1 ratio. And this rather odd-sounding finding was not a one-time fluke. In fact, this phenomenon has been found in many studies since then (Repp, London, & Keller, 2012). Somewhat paradoxically, the apparent attracting effect of the 1:1 ratio on the intended 2:1 rhythm causes people consistently to tap rhythms with a more complex ratio than they intend. Thus, although the serial ratio metric is related to the complexity of a rhythm, it may not provide a fully accurate measure.

Another reason the serial ratio can only be considered an approximation of rhythmic complexity comes from perception. Listeners do not perceive rhythms as precisely as the serial ratio concept implies. Take a ratio like 1:1.00003. Would a listener really hear that rhythm as complex, even though the second inter-onset interval would be just barely longer than you'd expect for a simple 1:1 ratio? Certainly not! Listeners only have a limited ability to discriminate different physical stimuli, as we discussed in chapter 3.

But there's an even more important phenomenon at play, which is commonly referred to as ***categorical perception***. This is a tendency to treat a range of values along a physical continuum as if they were the same until one reaches a point at which the percept abruptly changes. Categorical perception is a classic phenomenon in speech (for a review, see Jusczyk, 1997). Take the speech sounds 'pah' and 'bah.' You could ask five friends to produce these two syllables, and in each case you would hear the distinction between /p/ and /b/ as if your friends were producing the same two sounds. However, in all likelihood, each friend would produce /p/ somewhat differently from all the others, and likewise for /b/. Yet our perceptual system is insensitive to these small levels of variability because they are not important for perceiving the meaning of speech. Instead, we categorize these sounds as either being a /p/ or /b/, despite the minute acoustical differences that actually exist between variants of the two sounds.

We may also perceive musical rhythms categorically. Eric Clarke (1987) performed a series of experiments to test whether listeners have categorical perception for rhythm. Participants were asked to listen to 10 short musical items of five or six notes, regularly timed. Following this initial context, three test notes were played that formed a range of serial ratios between 1:1 and 1:2. Clarke found that people consistently heard the final pair of inter-onset intervals as forming either a 1:1 or a 1:2 ratio even when their timing deviated from these simple ratios. Moreover, listeners perceived the timing of these intervals as if they fit into two discrete categories separated by a boundary, in contrast to the physical reality of the stimuli. In a subsequent test, listeners had difficulty discriminating pairs of inter-onset intervals that did not cross the perceptual 'boundary' between 1:1 and 1:2. Subsequent research has extended this categorical perception to situations in which listeners classify more than two kinds of rhythmic categories (Honing, 2013).

Clarke proposed that the listeners discounted ratios that deviated from 1:1 and 1:2 because they were interpreted as the result of 'expressive information, or perhaps accidental inaccuracy' (1987, p. 30). Indeed, this observation converges with the way music is actually performed.

If one were to analyze performed ratios in recorded music, one would find few simple serial ratios, yet the music can sound quite 'rhythmic.' Chapter 11 (on music performance) will discuss why it is that performers might produce rhythms in this way. Categorical perception demonstrates that listeners can accommodate such variability.

Broader significance of rhythm

Although it is important to know how we perceive rhythmic patterns, one may be inclined to wonder whether the ability to perceive rhythmic patterns is relevant beyond the context of perceiving the temporal qualities of music. The answer to date is a resounding 'yes.'

First, the temporal qualities of music have important implications for how well we perceive *non*-temporal qualities of music, particularly pitch. A series of experiments carried out by Mari Riess Jones and colleagues explored the way in which rhythmic regularity allows the listener to recognize the pitch of a forthcoming note (Jones, Moynihan, MacKenzie, & Puente, 2002). They used a memory paradigm introduced by Deutsch (1972). The listener is presented with a 'standard' tone to hold in working memory while a series of subsequent 'distracter' tones are presented. Following the distracter tones, the listener hears a final 'comparison' tone and determines whether the pitch of the comparison tone matches or does not match the standard tone. Even though the distracter tones are irrelevant to the task, listeners used the timing of the distracter tones to form expectations for when the comparison tone would occur. As such, if the comparison tone occurred at a time that was consistent with the timing of the distracter tones, recognition of its pitch was better, particularly if the tones did not form a potentially distracting tonal pitch pattern (Prince, Schmuckler, & Thompson, 2009). Thus, rhythm is not just for rhythm. Listeners use rhythm to target their attention to forthcoming points in time (Jones, 1976; Jones & Boltz, 1989; Large & Jones, 1999).

Second, rhythm perception abilities may facilitate the perception of rhythms in nonmusical patterns. A 1997 study by Mangione and Nieman which appeared in the *Journal of the American Medical Association* examined over 500 physicians-in-training and medical students on their ability to identify common irregularities in recordings of human heartbeats. An unexpected secondary finding was that doctors who played a musical instrument were more accurate at identifying cardiac events than doctors who had no musical training. Cardiac auscultation is one of the most difficult diagnostic skills that doctors must master. 'In 0.8 seconds, you have four or five acoustic events at the threshold of audibility. You need to be able to separate them, and pick them up as a pattern,' explained Dr. Mangione when interviewed about the study (*The New York Times*, September 3, 1997). While a direct causal relationship cannot be drawn, it is possible that musical training hones one's ability to hear a rhythmic pattern against an imagined pulse and to detect deviations from it.

Third, the ability to perceive musical rhythms may even facilitate language use. Gordon and colleagues assessed how well children aged 5–7 years could discriminate rhythmic patterns, and compared that ability to the way in which these children answered questions about pictures they viewed (Gordon, Shivers, et al., 2015). Children who performed well on the rhythm discrimination task were also more competent at expressive grammar and used more complex syntax than children who performed more poorly. This facilitation may occur because of the way in which rhythm perception focuses attention, as discussed earlier. Specifically, in both music and language, sensitivity to rhythm may help the listener to focus on important forthcoming events and thus become better at processing information and generating plans for the verbal expression of one's thoughts (Gordon, Jacobs, Schuele, & McAuley, 2015).

Tempo

Tempo is the rate or speed of music. Conceptually, rhythm and tempo are independent of each other. Whereas rhythm refers to relative time, tempo refers to absolute time. The ability to perceive how fast or slow a rhythmic pattern is progressing may seem simple at first, and it would be if all rhythms were ***isochronous*** (meaning 'of equal time,' comprising only a single, repeating inter-onset interval). But as the previous section explained, rhythmic patterns are typically variable, and consist of various longer and shorter inter-onset intervals. So how does a listener get a sense for the speed or pace of the music?

Surprisingly few studies have addressed the process by which listeners perceive tempo in a complex rhythmical pattern. Repp (1994a) asked listeners to match metronome markings to expressive performances of Schuman's *Träumerei*. Although listeners tended to underestimate the tempo slightly, their ratings reflected the average of the inter-onset intervals in the performance. This suggests that listeners may be sensitive to the statistical properties of timing, similar to the way in which listeners are sensitive to the frequency with which pitches occur (another statistical property) when perceiving tonality (chapter 5).

Tempo and the 'beat'

In Repp's aforementioned study, participants' judgments were compared to notated metronome markings used to indicate the rate of the ***beat***, or ***tactus***: a salient isochronous time-span associated with our sympathetic movements to music (e.g., foot tapping). Is tempo, then, equivalent to the beat? It can be, but only up to a point. The beat is a distinct perceptual phenomenon that is a by-product of meter (which we discuss in the next section). In particular, there are constraints on perceived beat that are distinct from tempo perception. In general, people typically try to maintain the percept of a beat at a preferred rate that is consistent with the comfortable rate at which one can move periodically (e.g., walking, bobbing one's head). Early estimates of this rate placed it at 600 milliseconds (about 100 beats per minute) on average (Fraisse, 1982; Parncutt, 1994), although subsequent studies place it closer to 500 milliseconds (120 beats per minute; Moelants, 2002).

Consider a case in which a melody is repeated and is continually sped up after each repetition. During the first few repetitions, quarter notes (crotchets) occur every 500 milliseconds and establish the perceived beat. As the music speeds up, the perceived beat speeds up with each repetition for a while. But we will eventually get to a point where the durations of half notes (minims) are near 500 ms. Now the listener may start hearing the beat as being 500 ms again, but this will be based on the timing of half notes, which are now occurring at the quarter notes' original rate. Thus, the perceived beat can only speed up so much. Note, however, that when such a perceptual readjustment occurs, the listener is not fooled into thinking that the music has slowed down. In fact, the music can sound quite fast even when the beat may 'slow down,' as in this example. Similarly, when a Dixieland Jazz group enters 'double time' (literally doubling the speed), a listener is likely to keep tapping his or her foot at the same rate, but will easily hear the change in tempo. Thus, though beat and tempo are related, their relationship is complex, and not always equivalent.

More recent research casts further doubt on how closely aligned the beat is to tempo. Justin London (2011) compared listener judgments of how 'fast' or 'slow' the tempo of a rhythmic pattern is with identification of the 'beat' for these same patterns. He developed rhythmic patterns that had a constant beat, but varied in the rapidity of the rhythmic pattern that was demarcated by the beat (similar to the 'double time' example given before). In general, tempo judgments were predicted by the rate of the rhythmic pattern, rather than the rate of the beat, and were not strongly influenced by whether listeners attended to the beat. Part of the reason for this dissociation is that musical meter can be used to establish not one beat, but several possible beats that a listener can hear (Repp, 2011). Madison (2009) capitalized on this ambiguity and created an auditory pattern that gives the illusion of an ever-increasing tempo that never gets faster in the end – an auditory analogy to the Escher staircase (audio files are available by accessing the online research article at http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0008151, which is freely available to all readers).

Spontaneous tempo

Thus, tempo is not identical to the beat, yet it is true that the beat we hear has a rate that is related to both rhythm and tempo. The kind of tempo we associate with the beat is a question of interest in its own right, particularly the fact (mentioned before) that people are typically inclined to perceive a beat whose period is around 500 ms. Why is this? One possibility is that the beat, more so than perceived tempo, emerges from the rate of preferred periodic movements, like tapping one's foot. Specifically, the preferred rate of the beat is related to one's ***spontaneous tempo***, which was first explored by Paul Fraisse. Fraisse (1982) found that individuals are surprisingly consistent when asked to simply tap at a comfortable rate, which is how spontaneous tempo is defined. There is some variability in spontaneous tempo across individuals, but a surprising consistency in spontaneous tempo within an individual.

As mentioned before, the dominant preferred rate across individuals for beat perception and tapping appears to be around 500 ms (a bit faster than Fraisse's initial estimates). Fraisse also found that when people were asked to tap prototypically 'long' and 'short' durations, they tended to produce durations that were related by a 2:1 ratio. Hence, there is a possible motoric source for preferred serial ratios in rhythmic patterns, as discussed earlier. One of the factors influencing the preferred spontaneous tempo of the beat is age. Older adults generate slower spontaneous tempos than younger adults and children (McAuley, Jones, Johnston, & Miller, 2006). Similarly, adults synchronize to a slower beat than children when listening to music (Drake, Jones, & Baruch, 2000; see also chapter 9 in this volume). Musicians synchronize to a slower beat than nonmusicians (Drake, Penel, & Bigand, 2000), possibly because musicians focus on broader spans of time when listening to music than nonmusicians. Finally, spontaneous tempo is probably influenced in part by the kind of movement that is involved; for instance, spontaneous rates for piano performances are substantially faster than Fraisse's original estimates for tapping, and may be closer to 400 ms per IOI (Zamm, Wellmann, & Palmer, 2016).

Tempo, rhythm, and memory

We remarked earlier that tempo and rhythm are conceptually independent, but to what degree are they perceived in this way? For a certain range of tempos, independence does hold true. As mentioned before, you can speed up or slow down 'Shave and a hair cut …' and the distinctive rhythm will still be recognizable. However, the independence breaks down at extremely fast and extremely slow tempos, where the pattern of inter-onset intervals ceases to sound rhythmic at all (as may be the case when throwing pots down a staircase). Various research findings have led some to speculate that the perception of rhythms only holds for time-spans ranging from 200 ms (300 beats per minute) to 1000 ms (60 beats per minute; e.g., Drake & Botte, 1993). When inter-onset intervals are shorter than 200 ms, the notes are heard as too undifferentiated to create a rhythm (e.g., grace notes), or they may be heard as

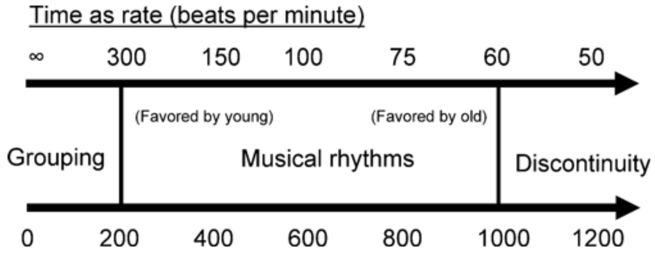


Figure 6.2 The continuum of musical tempos, delineating boundaries in which note events are typically heard as rhythmic.

subdivisions of rhythms (London, 2004). In contrast, intervals longer than 1000 ms are heard as separate events, and thus are not heard as being united within a rhythm. The exception is when the interval is subdivided by other tones, as when a solo instrument holds a suspended tone that is subdivided by accompanying instruments. Figure 6.2 illustrates the relationship between tempo and rhythm, generalizing across several studies.

Listeners have exhibited absolute memory for the tempo of familiar songs (Levitin & Cook, 1996). In some cases, this absolute memory for tempo may be life-saving! When reviving an individual's heart through cardiopulmonary resuscitation (CPR), it is critical to apply pressure with the appropriate rate. As it turns out, an ideal tempo to use for this procedure (100 beats per minute) is the tempo of 'Stayin' Alive' by the Bee Gees. In fact, the American Heart Association recommends this technique for use with hands-only CPR. Were it not for our keen ability to remember tempo, thanks to the song's title and popularity serving as a memory aid, fewer lives might be saved. There is a cautionary note here, however. Recent research suggests that our memory for tempo can be influenced by the tempo of a song we've heard very recently (Rashotte & Wedell, 2014). Thus, you may be want to be sure to 'clear your mind' of recently heard music before engaging in this procedure.

Meter

You have likely noted the prevailing importance of 'the beat' to musical time. However, we have not explained how the beat arises from variable rhythmic patterns. Indeed, there is a paradox in musical time concerning the fact that rhythms are variable, yet listeners spontaneously extract a regular period, and time rhythmic movements to that period. For most of us, this process is easy, although a case like Mathieu demonstrates that it may not be as simple as it seems. In this section, we introduce a manifestation of musical time that embodies the regularity from which the beat emerges: meter.

Defining and representing meter

Some readers will know of meter based on music notation, as a designation of a 'time signature' that establishes where 'bar lines' occur. The ***time signature*** is typically represented as two vertically oriented numbers near the left side of notation (see Figure 6.3), with the lower number indicating the kind of noted duration assumed to be a beat (often 4 for a quarter note or crotchet), and the upper number indicating the number of these beats that occur in each measure (also 4, in Figure 6.3). This manner of notating meter yields two important concepts. First is the ***beat***, the rate at which you find yourself bobbing your head

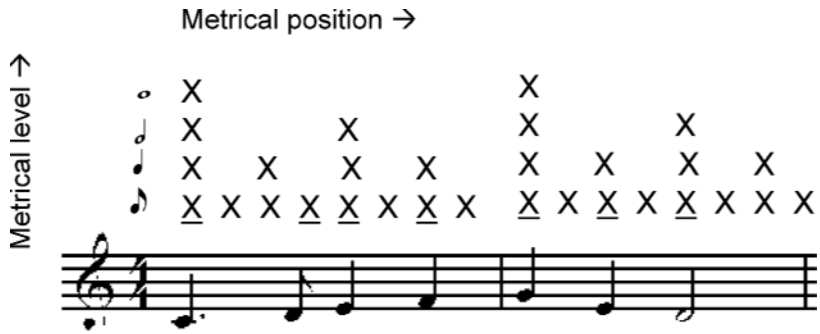


Figure 6.3 A metrical grid, shown as columns of X's, superimposed above a notated rhythm. Positions within the grid that are marked by note onsets are indicated by underlining at the lowest metrical level.

or tapping your feet as you hear a rhythm. Second, these beats are organized in recurring cycles that form a period. When the upper number is 4 (as in Figure 6.3), it means that the meter establishes a recurring four-beat pattern. Western music favors cycles that are multiples of twos (marches) or threes (waltzes), Indonesian Gamelan favors fours, while Indian music makes use of more complex meters based on higher prime numbers.

The definition of meter we just presented, based on its use in music notation, tells only part of the story. Music theorists have identified another important feature of meter that is not so apparent from its representation in music notation, and this leads to a fuller definition of meter that is the standard for psychology of music. ***Meter***, by this definition, is a repeating pattern of alternating strong and weak time points (including stronger and weaker beats). This alternation is formalized in the ***metrical grid*** shown above the notation in Figure 6.3 (i.e., the columns marked by X's). In this representation, the number of X's at a specific point in time determines how prominent or ***accented*** that time point is, with more X's at a given point indicating greater prominence. Note that the grid is cyclical, in that X's at each level repeat regularly. The grid here matches the counting of 4/4 time, but also establishes which sounded notes are more prominent (e.g., the first note and to a lesser extent the third note of the notated melody) and which are less prominent (e.g., the second note and fourth notes of the melody).

The metrical grid notation gives us a deeper understanding of the beat in music. We mentioned before that the lower number in a time signature determines the beat. However, in reality, the perception of what constitutes the beat is up to the listener, and is more flexible than this. Notice that the left-most column in the grid shows four different notes, each with a different duration value. Together they form a hierarchy of time-spans, with each duration at a higher level coming from the sum of two or more durations at a lower level. Each level represents a time-span that could be heard as the beat. For the example in Figure 6.3, most people would hear the beat at the quarter note (crotchet) level, hence the time signature. However, it would be possible to hear the beat at other metrical levels (e.g., the half note, or minim level), particularly if the rhythm is performed at a fast tempo, as per the previous section. Typically, the note value lying somewhere in the middle of this metrical hierarchy establishes the beat. Levels above (of longer duration than) determine which beats are accented, whereas levels below the beat subdivide the beat.

In our culture, 'simple' meters are common. ***Simple meters*** are those in which the meter's cycle can be reduced to a count of two or three. This includes meters with time signatures like 4/4 and 6/8. They are referred to as simple because there is an isochronous (consistent) duration separating all of the most strongly accented beats (e.g., London, 1995). Meters with a count that is not reducible to two or three are called ***complex meters***. Complex meters often highlight periods of five, seven, or nine. Two of the best-known examples include the jazz piece 'Take 5' (a count of five) and the Pink Floyd song 'Money' (a count of seven). In cultures where complex meters are rarely found (like much of Western Europe and North America), listeners may find the associated rhythm pattern unusual and not suitable for dancing. However, there are cultures, such as the Balkans, in which complex meters are common and frequently used for dancing.

An important set of studies by Erin Hannon and Sandra Trehub (2005a) demonstrated that adults who had never been exposed to Balkan music had difficulty noticing perturbations of rhythms based on a complex Balkan meter, whereas listeners raised in the Balkans had no trouble with this task. At the same time, infants from non-Balkan households were able to perform similarly to the Balkan listeners after only two weeks of exposure to Balkan meters (Hannon & Trehub, 2005b; see also chapter 8 in this volume). Thus, the dominance of simple meters in many cultures may reflect their traditions rather than an innate predisposition for listeners to prefer simple meters.

The relationship between meter and rhythm

Figure 6.3 illustrates this match between meter, represented by the metrical grid, and rhythm, represented by the patterning of note onsets. Why are meter and rhythm thought to be different manifestations of musical time? One reason has to do with reliance on sound itself (the musical surface; see chapters 5 and 7). Rhythm relies on the presence of event onsets. Metrical accents, such as the beat, can exist when no onset is there. In some cases, a metrical accent can coincide with a silent point in time; as a result, musical silences can seem conceptually 'loud' (Margulis, 2007). A classic example that Margulis cites occurs in Beethoven's *Eroica* Symphony No. 3. At one point in the first movement (measure 280), an unexpected silence appears at the beginning of a measure – a strongly 'accented' silence.

A second reason for distinguishing rhythm from meter is that they need not agree, and much of our enjoyment of music has to do with their disagreements. An important example is ***syncopation***, which is pervasive in genres like rock and jazz. Syncopation occurs when event onsets align with weak metrical accents, yet are conspicuously absent at strong accents. Paradoxically, syncopation does not necessarily cause the underlying beat to become ambiguous. David Temperley (2000) has pointed out that in rock music the sung lyrics regularly occur just prior to a strong metrical accent, and are thus dominantly syncopated in a way that anticipates the beat. 'Let It Be' by The Beatles is a good example. The accented syllables (shown in capital letters) in 'MO-ther MA-ry COMES to ME' all anticipate the regularly timed piano chords. He argues that the use of syncopation in this way draws the listener's attention to the beat, rendering the music more danceable. In general, music that promotes motion, usually referred to as ***groove*** (think of virtually any song by Stevie Wonder), is moderately syncopated (Ashley, 2014; Iyer, 2002; Janata, Tomic, & Haberman, 2012; Witek, Clarke, Wallentin, Kringelbach, & Vuust, 2014).

If meter and rhythm are truly separate, where does meter come from? Meter results from the intersection of information from incoming rhythmic patterns with learned schemas. It is likely that listeners base their interpretation of meter on the first rhythmic patterns in a piece (Longuet-Higgins & Lee, 1982). Listeners fit this pattern into a pre-existing metrical schema stored in long-term memory. As the music continues, listeners maintain their initial interpretation as long as it is reasonably consistent with unfolding rhythmic patterns. Thus, in most cases, rhythm leads to the interpretation of a meter.

However, it is possible to reverse this chain of events such that a metrical interpretation influences how a rhythm is perceived. For instance, Palmer and Krumhansl (1990) asked listeners to imagine a particular meter by counting silently, and then presented listeners with simple tones at various times. They found that listeners thought tones sounded ‘better’ when tones coincided with strongly accented positions, based on the meter that the listener was imagining. This finding was particularly strong for musically trained listeners. In a similar vein, Desain and Honing (2003, experiment 2), found that a preceding metrical context (a beat pattern), influenced how musicians categorized ambiguous rhythms presented afterwards. Meter also persists when listeners hear a song that ‘fades out,’ which makes the ending point of a song ambiguous. Unlike songs that end definitively, fade-outs cause listeners to continue to hear a metrical pattern after the auditory signal has ceased (Kopiez, Platz, Müller, & Wolf, 2015).

Several papers have examined the way in which the fit between rhythm and meter can influence the perception of rhythmic complexity and the perception of a ‘beat.’ Perhaps the most influential study in this group is a paper by Povel and Essens (1985). The authors proposed a model based on finding the best match between a rhythm and a time period for the beat. A good match happens when beats co-occur with tone onsets rather than silences. Rhythms that are simpler and more ‘patterned’ yield a clear match, whereas more complex rhythms, including those that do not sound patterned, do not clearly match a specific beat. In general terms, this model suggests that our perception of rhythm is influenced by the degree to which a rhythm can be heard as belonging to a particular meter.

An illustration of the approach advanced by Povel and Essens is illustrated in Figure 6.4. Two rhythms are represented by vertical lines (tone onsets) superimposed above rows of dots that represent candidate time points for tone onsets. Note that both melodies contain simple serial ratios. Thus the perceived complexity of rhythms would only result from how well each rhythm matches a particular beat.[3] As can be seen, the top rhythm clearly matches a beat that groups rhythms into patterns of four (using the dots as the base unit), in that tone onsets always align with metrical boundaries (the beat). However, counting in threes (as in 3/4 meter; see the brackets below the rhythm) does not work as well for this rhythm. By contrast, the lower (complex) rhythm, does not match either organization very well, as metrical boundaries for both organizations can fall on silent points. Povel and Essens found that rhythms like the one shown at the bottom of Figure 6.4 were more difficult for listeners to remember and reproduce than rhythms like the one shown above it.

The neural bases of rhythm perception

Rhythms in the brain

In considering the role of rhythm in the brain, one might consider the fact that brain activity is inherently rhythmical. As we discussed in chapter 4, EEG measurements show rhythmic oscillations in neural activity during different cognitive states. The fact that such neural rhythms exist, however, does not mean that these neural rhythms play a role in the perception of musical rhythms. But it is a possibility that has been pursued at the theoretical, behavioral, and cognitive neuroscience levels. A dominant theoretical notion in the literature is that the brain engages in ***entrainment***, which is a process in which one rhythmic pattern achieves and maintains synchrony with another pattern (Large & Jones, 1999). According to this view, neural rhythms that are intrinsic to the brain (sometimes called ‘endogenous’) synchronize

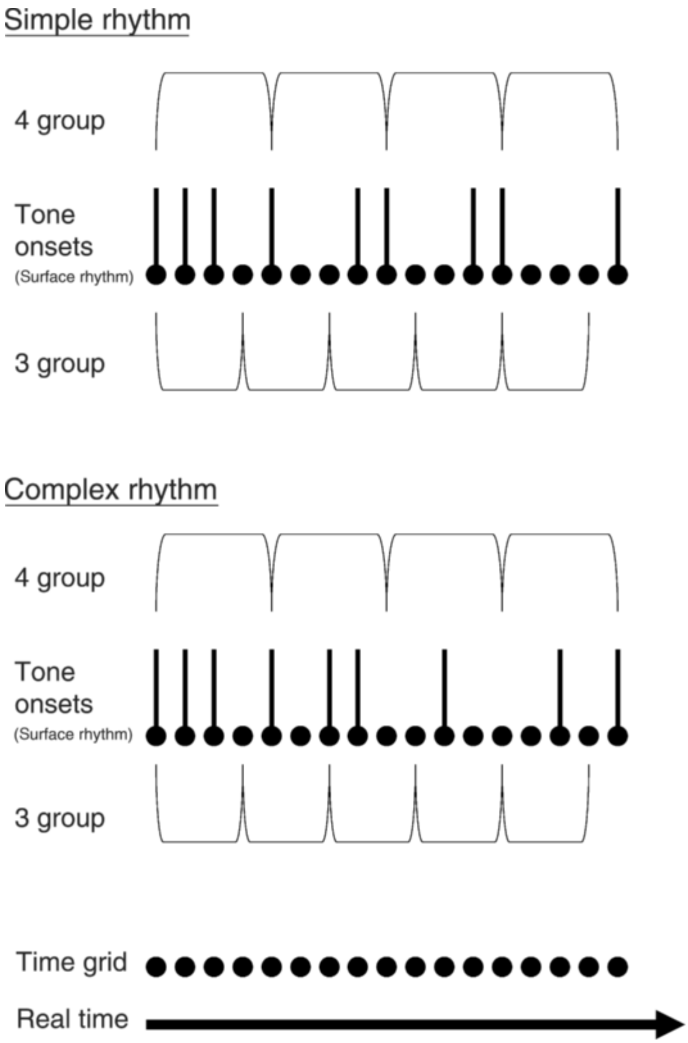


Figure 6.4 Two examples of rhythms used by Povel and Essens (1985). Time in this approach is discrete, represented by dots for candidate time points. The two manifested rhythms (simple, complex) are shown as vertical lines representing tone onsets above dots; dots presented alone represent silence. Brackets above and below represent candidate organizations based on beats that recur every four time points (above) or every three time points (below).

Source: Adapted from Povel and Essens (1985, Figure 5, p. 420) with permission from the University of California Press.

with external rhythms present in music, much like a dancer synchronizes body movements to music. More recently, the Action Simulation for Action Prediction (ASAP) hypothesis has extended the role of entrainment beyond perception (Patel & Iversen, 2014). According to this view, entrainment to rhythms engages motor areas of the brain, and this motor activity allows us to predict forthcoming events. It is this prediction that allows us to dance or be disrupted in a case like Mathieu, described at the start of this chapter.

The strongest support for entrainment in the brain comes from measures of neural dynamics, namely EEG/ERP (see chapter 4). A particularly critical test is whether neural activity anticipates where tone onsets ought to be, or where accented tones ought to be, based on extracting a sense of beat or meter. One of the earlier studies on this topic used the brain's 'surprise' response in ERPs to determine whether the brain anticipates where accents should be (Brochard, Abecasis, Potter, Ragot, & Drake, 2003). The authors took advantage of a well-known auditory illusion in which listeners hear alternating tones as loud and soft when presented with a sequence of tones that are in fact physically identical (having the same pitch, duration, and loudness). This phenomenon is known as ***subjective accenting*** (Bolton, 1894).

Brochard and colleagues found that the brain's 'startle' response followed a typical pattern of accenting on every other tone (as in a binary meter). When tones on subjectively 'strong' accents were made suddenly quiet, the ERP response showed greater 'surprise' than when tones on subjectively 'weak' accents were made suddenly quiet. More recent papers have addressed neural entrainment based on meter. To do this, one has to look at more rapidly changing aspects of the EEG signal than are found in ERP studies. Some evidence suggests that rapid oscillations in the brain can anticipate tone onsets, and even respond to silences at which tone onsets are expected (Fujioka, Trainor, Large, & Ross, 2009). In chapter 8, we discuss how neural responses to rhythm can be found even among newborn babies!

Where might such entrainment occur? In chapter 4, we discussed how the cerebellum plays a role in timing. This is true, but it appears as though the cerebellum is more important for keeping track of absolute time, rather than relative timing that is critical for rhythm (Teki, Grube, Kumar, & Griffiths, 2011). Some evidence suggests that the cerebellum plays a role in processing complexity of rhythms' serial ratios (Bengtsson et al., 2009; Sakai et al., 1999). However, statistical support for this role has been weak (Grahn, 2012). Indeed, if there has to be one most critical area for neural entrainment, it is likely to be the basal ganglia.

In an influential paper, Grahn and Rowe (2009) analyzed brain areas using fMRI when listeners assessed how well a regular pulse (presented as a series of tones) matched a musical rhythm. Many areas of the brain showed different levels of activation based on how well rhythms matched the pulses that were presented. However, when the authors went on to analyze statistically which areas of the brain would most strongly predict activation in other areas, they found that the basal ganglia stood out as a critical predictor. Clinical evidence offers further support. One of the devastating effects of Parkinson's disease is a breakdown of rhythmic motor behavior in activities like walking (a rhythmic, though not musical, behavior). These deficits occur because Parkinson's involves disruption of the basal ganglia, which is responsible for producing the neurotransmitter ***dopamine***, and dopamine is critical for controlling voluntary movements. Although walking is a nonmusical behavior, it is rhythmic, and so synchronization with musical rhythms can facilitate the re-learning of walking through stimulation of the basal ganglia (Nombela, Hughes, Owen, & Grahn, 2013).

Neural entrainment is presumably linked to our physical responses to musical rhythms. We mentioned such responses in our previous discussion of groove and the ASAP hypothesis. The neural basis for this close link between musical rhythm and movement was revealed in further analyses by Grahn and Rowe (2009). In addition to the basal ganglia, the authors found activation in the supplementary motor area, which typically plays a major role in motor planning (such as planning one's dance moves). Other research has found activation in a nearby region, the premotor cortex, when participants hear rhythms that are clearly metrical (e.g., Chen, Penhune, & Zatorre, 2008). The significance of this research is that the premotor cortex is more tightly coupled to overt actions than the supplementary motor area, yet there are areas within the premotor cortex that are active even when listeners have no intention of moving to the beat.

Auditory specialization for rhythm?

Of course, rhythms in music are part of the overall auditory signal. This brings up two related questions: Does the auditory cortex show any specialization for rhythm? And to what degree do different parts of the brain contribute to rhythm versus pitch perception? With respect to the first question, there is some evidence that the left temporal lobe, more so than the right temporal lobe, is necessary for rhythm perception, although this specialization may be limited to very short time intervals (Samson & Ehrlé, 2003). Patients who have had portions of their left temporal lobe removed in order to prevent seizures have difficulty discriminating short time intervals that cause perturbations of rhythmic sequences. With respect to the second question, behavioral evidence gives a mixed picture, with some studies suggesting that melody and rhythm are processed independently (e.g., Palmer & Krumhansl, 1987), and others suggesting integration (e.g., Boltz & Jones, 1986).

As one might expect, evidence from neuroscience is similarly mixed. Studies of patients with brain damage suggest that different brain regions are necessary to process pitch and time. Cases have been reported in which brain damage to the temporal lobes interferes with pitch discrimination, but not time perception (e.g., Liégois-Chauvel et al., 1998), as well as damage that hinders rhythm discrimination, but not pitch perception (e.g., Peretz, 1990). These neural dissociations can account for someone like Mathieu (discussed at the beginning of the chapter), who did not suffer brain damage, but may have had poor functioning in brain areas necessary to process musical time. At the same time, an fMRI study found overlapping brain regions, primarily within the right auditory cortex, when listeners tried to detect changes to pitch structure or rhythm in melodies (Griffiths, Johnsrude, Dean, & Green, 1999). In fact, the only difference between these conditions was a region of activation in the dorsal cerebellum that occurred during rhythm detection, but not pitch detection. Yet it is not clear how exclusively temporal the cerebellum is, given other evidence suggesting that the cerebellum may be active during pitch perception (Petacchi, Laird, Fox, & Bower, 2005). Overall, the mixed findings from both behavioral and neuroscientific research suggest that the interplay between melody and rhythm is still not fully understood.

Coda

Rhythms are intrinsic to the way we move and interact in the world, and are inarguably a deep part of musical experience. Rhythms shape time (Epstein, 1995). Though time is commonly thought to be one-dimensional, like an arrow, rhythmic structure alters this state, leading to a temporal continuum rich with cyclical and hierarchical structures. Theorists have argued that different forms of temporal organization run concurrently as music unfolds, creating sources of tension at times that make music more pleasing and interesting when one is listening to it. The basis for our engagement with musical rhythms may be due to the close link between rhythm and movement, at a neural level. As such, listening to rhythms brings about a kind of internal 'dance' between the listener and the music, characterized by figural patterns (rhythm), speed (tempo), and an underlying cyclical structure (meter).

Notes

1 The first sentence here is also the title of a scientific paper written for young audiences by Phillips-Silver (2014).
2 We are indebted to Caroline Palmer for this example.
3 There is a second component in their approach that we have left out for the sake of simplicity.

7 Analysis and cognition of musical structure

Many a curious child has taken apart a mechanical toy to see what parts are inside it, how they interlock, and what makes them work. Listeners often do the same with pieces of music. Careful listening involves determining how a piece of music is constructed, how the parts fit together, and what makes them work as a whole. It is easy to overlook the complexity of this process. Analyzing a thing in motion, with all its intricate dynamic properties, poses a formidable task. It takes a trained eye to identify patterns in the movements of a gazelle bounding through a forest, or a flock of birds in flight. Similar challenges are true of listening, as the listener hears a succession of tones that each only exists for an instant in time as a composition is unfolding. Furthermore, the systems that regulate perception and memory are limited with respect to the amount and type of information that can be processed (Snyder, 2000).

Although 'analysis' is often equated with musical expertise, it is important to note that any listener, to some degree, can take part in 'analytical' listening. Although not all listeners are musically trained, most at least grasp sufficient basic patterns or relationships to feel an overall sense of coherence in a new piece of music, to form general expectations about what might come next, to feel somewhat puzzled or frustrated when the music strays dramatically from what they expected, to sense when a phrase (or a whole piece) may be nearing the end, and whether the phrase or composition comes to a satisfying close or not. All the while, they are also organizing the incoming tones into some sort of meaningful array and into coherent units, such as melody and accompaniment, and separate phrases. This process may occur implicitly (see chapter 5), without conscious awareness, but implicit or not, our understanding of musical structure is integral to our experience of music.

Musicians – and particularly musicologists (music theorists) – develop the listening skills and vocabulary necessary to develop such analyses consciously and explicitly. Sometimes this process involves an examination of the written musical score, note by note, to more fully understand the interlocking parts – identifying the themes and motifs and their variations, the sources of unity, the harmonic underpinnings of a piece of music. This is the process that is typically referred to as ***analysis*** of music. Importantly, this explicit understanding of musical structure can lead to theories concerning how listeners come to comprehend musical structure while listening.

This chapter provides a general introduction to two of the most influential analytical approaches that have shaped our understanding of music analysis and cognition. These approaches are particularly interesting for the psychology of music, moreover, because they are predicated on an understanding of the psychological bases of listening, composing, and performing. It is worth noting that an analytical theory of music is not always a theory of how the listener or performer conceptualizes music. Traditionally, music analysis has been used to elucidate the structure of music, which may or may not be part of anyone's experience of that music. More recently, an increasing number of musicologists, led by the example of Leonard Meyer (1956), have become interested in the basic perceptual and memory processes that contribute to musical comprehension.

Lerdahl and Jackendoff's (1983) *Generative Theory of Tonal Music* concerns how a listener organizes aspects of the music that are hierarchically structured (such as motifs nested within phrases nested within sections of a composition, or beats and measures set within a metrical structure). Similar to the earlier approach of Heinrich Schenker (1935), described briefly later in this chapter, Lerdahl and Jackendoff's theory seeks to 'strip away' all the nonessential notes to uncover the very core of a musical piece. Their theory identifies four components of music that are hierarchically structured, and identifies a system of 'rules' that listeners intuitively apply in order to understand the piece. Even though listening itself is a dynamic activity, their theory is concerned only with the 'final state' of the listener's understanding (1983, p. 4). It seeks to capture the observer's lasting impression after the gazelle has disappeared into the woods, after the flock of birds has vanished into the horizon.

Narmour's (1990, 1992) *Implication–Realization model* departs from the hierarchical approach of Lerdahl and Jackendoff, and focuses instead on the musical surface. This theory, which finds its roots in the work of Leonard Meyer (1956; see chapter 13 for further discussion), is concerned with how a listener's expectations are formed – for instance, by what two successive tones imply may come next, and how this implication is realized, to explain the analysis and cognition of basic melodic structures. As its name implies, this model seeks to capture both the prospective and retrospective aspects of perceiving, structuring, and comprehending melody. Thus it seeks to capture the listener's strategies in melodic cognition while the piece is in motion, along with listeners' expectations and some retrospective analysis.

Due to the technical nature of the topic, it should be noted that the present chapter is different from all others, as it is the only chapter that assumes previous knowledge of music theory on the part of the reader. Although every effort is made to explain musical terms and concepts in all other chapters, it would be an unwieldy task to do so in the present discussion.

The Generative Theory of Tonal Music

A central assumption of mainstream cognitive psychology is that people apply internalized rule systems to guide their interpretation of events. When it comes to the interpretation of organized sequences, like music or language, many have claimed that these rules take the form of implicitly learned *grammars*. Although we associate the term with language, a ***grammar*** can refer to any system or rules designed to organize a small number of basic units (notes, words, sounds) into a vast number of meaningful sequences. Probably the most famous proponent of this idea as applied to language was Noam Chomsky. Chomsky's famous generative model of the cognitive foundations of language competence (Chomsky, 1966) asks us to imagine that the sentence forms we hear or read are associated with abstract rule systems that can be used to generate any possible grammatically correct sentence (including transformations of the sentence you just heard or read).

Chomsky's view was fundamental in the formation of one of the theories we focus on here, Lerdahl and Jackendoff's (1983) Generative Theory of Tonal Music. For example, summing up the basic principles of cognitive psychology of musical composition, Lerdahl (1988, p. 233) suggests that in any process of composition, a 'compositional grammar' is involved, which could be represented by systems of rules. These rules may express the knowledge and skills of a composer who creates a sequence of musical events, or of a listener trying to make sense of the same sequence. As Lerdahl explains, 'the listening grammar then generates the mental representation that comprises the "heard structure" of the piece' (p. 234). In so doing, the listener makes use of whatever body of musical knowledge he or she has so far acquired.

Background: Schenkerian analysis

Before describing the Generative Theory of Tonal Music in more detail, it is important to describe a historically antecedent musical theory, that of music theorist Heinrich Schenker. Schenker (1954, p. 29) believed that the deep structures of music could be revealed by a method of 'reduction' (which we will describe shortly) to bring out the basic structures that grounded the music, conceived as an elaboration of these primitive forms. Though their work is thought to have been independently developed, there is a common spirit in Chomsky's and Schenker's work (see Sloboda, 1985, pp. 11–23). Both are interested in deriving abstract structures that underlie complex sequences.

According to Schenker, every recognizable melody has a deep structure, or *ursatz*. The *ursatz* is complex, having a melodic line (the *urlinie*) and a bass line (the *bassbrechung*). Schenker proposed that the *urlinie* is built from the descending sequence of third, second, and tonic (E D C in C major), whereas the *bassbrechung* is built on the tonic, dominant, tonic sequence (C G C in C major). Schenker suggested that these deep structures are natural and universal (see chapter 15 for further discussion of universal and culture-specific aspects of music). As such, the listener's capability of melody recognition may rely on the generic *ursatz*. The tonic triad (first, third, and fifth tones) pattern can also be in the repertoire of *ursatzen*, since it is clearly the basis on which the whole scheme is built. Melody, then, must consist in departures from and returns to the *ursatz* defined by the basic Schenkerian structures. This is the basis of anticipation, the most important psychological state for the listener's experience of music.

The ***reduction*** of a piece of music to reveal the Schenkerian deep structure involves several steps by which the deep structure is revealed as we strip away or 'reduce' the intervening notes that constitute elaborations of base structures. By proposing that all tonal music compositions are generated from a simple fundamental structure (to which they can subsequently be 'reduced'), Schenker provided a way of isolating the distinctive features of a composition that set it apart from features which characterize all tonal music. Figure 7.1 shows an example of a pitch reduction of the Bach chorale *O Haupt voll Blut und Wunden* in the spirit of Schenker, but taken from Lerdahl and Jackendoff (1983, p. 108). We should note, however, that the primary goal of Schenker's theory was to identify a universal *ursatz*, from which all melodies could be generated, a kind of 'top-down' approach. The use of Schenkerian theory to create 'reductions,' as we have described, is closer to the goal of most theories in music psychology, and constitutes a kind of 'bottom-up' approach (Lerdahl, 2009).

An overview of GTTM

Similar to the ideas of Schenker and Chomsky, music composer and theorist Fred Lerdahl and linguist Ray Jackendoff (1983) set out to describe principles through which an underlying 'structural description' of music can be derived in their ***Generative Theory of Tonal Music (GTTM)***. Put simply, their theory aims to identify the specific rules by which a listener generates a 'heard structure' from the 'musical surface.' Lerdahl and Jackendoff (1983–1984)[1] define ***musical surface*** broadly as 'the physical signal of a piece when it is played' (p. 229) and ***heard structure*** as 'all the structure a listener unconsciously infers when he [or she] listens to and understands a piece, above and beyond the data of the physical signal'

Figure 7.1 Schenkerian-style reduction of the Bach chorale *O Haupt voll Blut und Wunden*.

Source: Lerdahl, Fred, and Ray S. Jackendoff, *A Generative Theory of Tonal Music*, Figure 5.4, p. 108, © 1983 Massachusetts Institute of Technology, by permission of The MIT Press.

(pp. 229–230). In the broadest sense, therefore, GTTM offers a 'musical grammar' in the form of an elaborate set of rules that explain how listeners come to hear sounds such as pitches of different durations and loudness as organized music. The applicability of the theory is restricted to tonal music, and the musical grammar focuses only on hierarchically organized aspects of music. The theory consists of four components, each devoted to one of four types of hierarchical structure (referred to as *Grouping Structure*, *Metrical Structure*, *time-span reduction*, and *prolongational reduction*), each subject to a set of rules.

GTTM's reductions (and in particular, prolongational reductions) are broadly similar to Schenker's in that an elaborate musical 'surface' is 'reduced' to its core constituents. However, the two theories differ in their aims. Lerdahl and Jackendoff view GTTM as a 'psychological' theory that aims 'to find principles of music cognition,' and Schenker's as an 'aesthetic' theory whose purpose is 'to illuminate musical masterpieces' (Lerdahl and Jackendoff, 1983–1984, pp. 248–249). The theories also differ in their focus. For instance, GTTM focuses far more on rhythmic elements than does Schenker's approach.

A musical grammar must identify the kinds of structures that are 'preferred' or 'not preferred,' and 'acceptable' or 'unacceptable,' within a given rule system. Thus, the rules associated with each of the components occupy an important role in GTTM. In Lerdahl and Jackendoff's (1983) words, 'each rule of musical grammar is intended to express a generalization about the organization that the listener attributes to the music he [or she] hears' (preface, p. x). In other words, their rule systems are designed to suggest a kind of 'structural description' that the listener forms while listening to music. There are two primary rule categories they propose. ***Well-formedness rules*** characterize a fairly fixed set of rules that distinguish 'acceptable' from 'unacceptable' structures. By contrast, ***preference rules***, which are more flexible, characterize the kinds of descriptions that a listener will gravitate to when more than one description is possible. In the original theory, there was some ambiguity concerning which preference rules would 'win' in specific situations. More recently, Temperley (2001) has developed an extension of the preference rule approach that features a quantitative rule system that can be used to determine which combination of preference rules best matches the interpretation of a listener.

A notational system is needed to visually represent the 'heard structure' of the music, and the general features of this symbol system are also shown in two reductions of Bach's *C Major Prelude* in Figures 7.2 and 7.3. (These figures are only provided as an illustration of the basic features of GTTM's notation system, and a detailed analysis of this composition can be found Lerdahl and Jackendoff, 1983, pp. 260–264.)

Grouping Structure

Grouping Structure 'describes the listener's segmentation of the music into units of various sizes' (Lerdahl and Jackendoff, 1983–1984, p. 231), each nested within the other (thus hierarchically structured), such as motifs within short phrases, within longer phrases, within sections of music, and finally, within the whole composition. In GTTM's notational style, these are represented by a series of slurs beneath the musical notation, with the broadest grouping (in which all smaller groups are nested) shown at the lowest layer, as shown in Figure 7.2. According to GTTM's formal grammar, listeners group the musical surface into segments or chunks consistent with two sets of rules ('well-formedness' and 'preference' rules) that are specified for each of the components.

Here, for example, are five 'well-formedness' rules for groupings of notes (Lerdahl and Jackendoff, 1983, pp. 37–39):

(1) Any contiguous sequence of pitch-events, drum beats, and so on can constitute a group, and only contiguous sequences can constitute a group.
(2) A piece of music constitutes a group.
(3) A group may contain smaller groups.
(4) If a larger group contains a smaller group, all the elements of the smaller group must be elements of the larger group.
(5) If a larger group contains a smaller group, the larger group must be exhaustively partitioned into smaller groups.

Though these principles may sound abstract, a simple example shows that they are in fact highly intuitive. Imagine the first line of the tune 'Twinkle, Twinkle, Little Star':

Twinkle, twinkle, little star, how I wonder what you are.

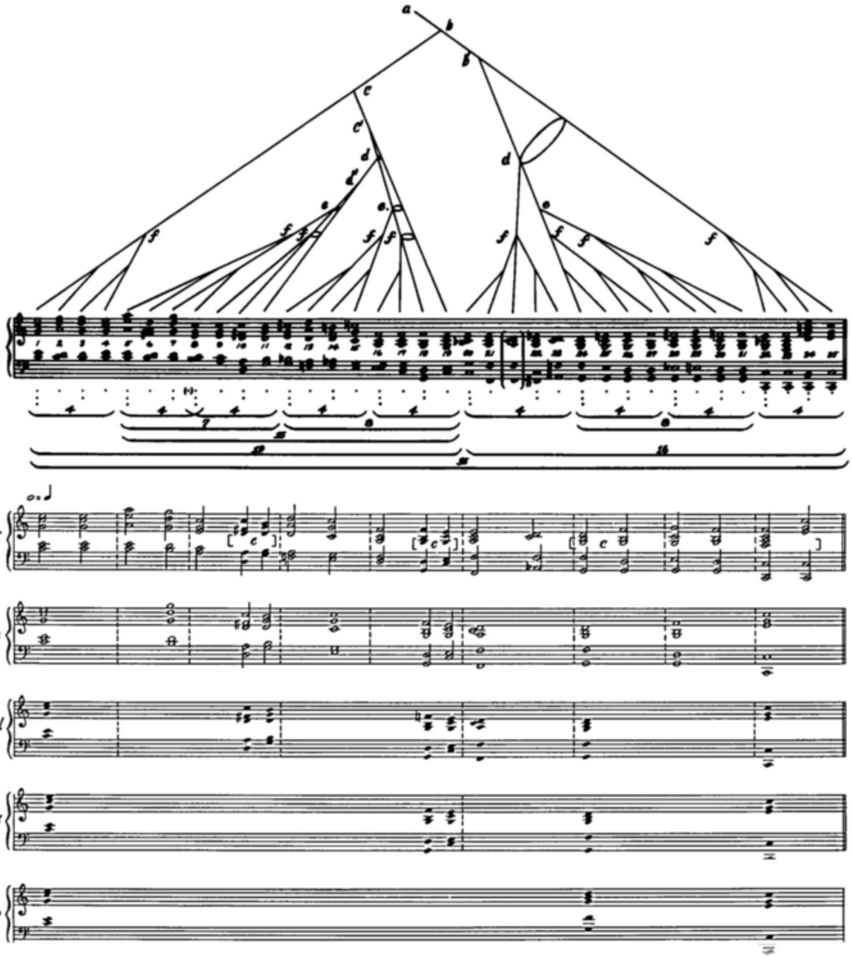


Figure 7.2 Time-span reduction of Bach's *C Major Prelude*.

Source: Lerdahl, Fred, and Ray S. Jackendoff, *A Generative Theory of Tonal Music*, Figure 10.10, p. 262, © 1983 Massachusetts Institute of Technology, by permission of The MIT Press.

This line when sung may be considered to be a single musical phrase, and thus a group. At the same time, we can hear two smaller groups within this phrase, one functioning as a kind of musical 'question' ('Twinkle, twinkle, little star') and the other functioning as an 'answer' ('how I wonder what you are'). Note that these two smaller groups are bounded by the larger group. It simply does not make sense conceptually to have a smaller group (a subphrase) cross a boundary that is formed by a larger group. Furthermore, this single phrase is one of several that together make up the entire tune, and the two subphrases discussed above can likewise be carved into yet smaller groups, the smallest of which would comprise individual notes.

Preference rules are greater in number and more complex than well-formedness rules, and so we will not enumerate all grouping preference rules here, but instead focus on two examples. Generally speaking, grouping preference rules suggest where a listener will typically, but not always, hear a group boundary. One such rule is based on ***temporal proximity***. As in Gestalt psychology (see chapter 5), this rule suggests that a relatively long time-span between notes will often be heard as a group boundary. Going back to our example, the lengthened note associated with the first singing of 'star,' in comparison to the shorter notes associated with 'twinkle, twinkle, little . . .,' would lead a listener to (most likely) hear a group boundary where the lengthening has occurred.

But proximity clearly is not sufficient. One could extend the rule of proximity to the point that every single note is its own group; after all, there is a 'boundary' after each note in a melody (disregarding chords). To guard against such counter-intuitive 'preferences,' Lerdahl and Jackendoff (1983) also propose a rule stating that listeners prefer groups that are not too short in length (certainly not just one event).

Metrical Structure

Other rules are based on a piece's ***Metrical Structure***, which constitutes the pattern of alternating strong and weak time points that typically recur in cycles that are based on multiples of two or three (see chapter 6 for a more detailed discussion). Importantly, metrical structures are hierarchical, such that strong accents recur at long time-spans that subsume the smaller time-spans between weaker accents and unaccented time points. This metrical hierarchy is represented by a grid of X's or dots, as illustrated in chapter 6.

The well-formedness rules for identifying metrical structure include:

(1) Every attack point must be associated with a time point at the lowest level of the metrical hierarchy.
(2) Every time point at a given level must also be found at all smaller levels.
(3) At each metrical level, strong accents are spaced either two or three time points apart.[2]
(4) Each metrical level must consist of equally spaced time points (Lerdahl & Jackendoff, 1983, p. 69).

Again, these seemingly abstract principles lead to more intuitive results, in that they lead to metrical structures that are both regular (with respect to spacing of time points) and alternating (with respect to the prominence of accents). For the listener, the result is that meters frame music with respect to alternately stronger and weaker points in time, an effect that we discuss in more detail in chapter 6. For instance, the melody 'Twinkle, Twinkle, Little Star' yields the strong sense of alternating strong and weak accents on different syllables, with the first syllable of each 'twinkle' being associated with a strong accent.

Preference rules for meter address the thorny issue of which events are interpreted as strong versus weak, as well as the link between these patterns of accenting on the musical surface and the underlying meter that a listener interprets from that music. The link from musical surface to meter is not always straightforward, as we discussed in chapter 6. As with grouping, GTTM suggests many metrical preference rules, from which we select two examples. Perhaps the simplest preference rule is referred to as the ***event rule***. It stipulates that strong metrical accents should coincide with the onset of notes. Thus a listener is unlikely to hear a strong beat in the middle of the sustained note for 'star.' Another, complementary rule suggests that stronger beats are associated with the onset of relatively longer notes. Thus, one is unlikely to hear a strong beat in the middle of 'star,' and is instead likely to hear a strong beat at the beginning of that note.

Time-span reduction

A ***time-span reduction*** 'establishes the relative structural importance of pitch-events within the heard rhythmic units of a piece' (Lerdahl & Jackendoff, 1983, p. 231). Formally, this is notated as a tree diagram that 'expresses the way in which pitch-events are heard in the context of hierarchically organized rhythmic units' (p. 237). In other words, these reductions allow us to identify the relatively important pitch events (tones) in each rhythmic group, which determines the underlying 'deep' pitch structure. Finding Schenker's notation 'not explicit enough' (Lerdahl & Jackendoff, 1983, p. 112), GTTM borrows the 'tree' notation of linguistics to express the hierarchical nature of reductions. However, they point out that the similarity with linguistic trees is cosmetic; theirs are 'purely musical trees, having nothing to do with linguistic trees except that both express hierarchical structures with precision' (p. 112). The tree notation appears above the musical notation, and can be quite elaborate for complex compositions. Underneath the musical notation, the reduction appears on a separate series of staves, as seen in the time-span reduction of Bach's *C Major Prelude* in Figure 7.2.

The basic procedure for notating a time-span reduction takes several steps. One starts with a hierarchy of rhythmic groups using principles sketched in the previous sections. In each time-span, for each level of the hierarchy, one identifies a single chord or pitch event that functions as 'the head' (i.e., is prominent) according to whether it acts as a structural point of resolution within that time-span. For instance, in the first group of four chords in Figure 7.2, the initial C-major chord functions as the head; this is evidenced by the fact that the 'branch' from this chord extends over the branches from all other chords, thereby encompassing the rest. As one moves up the hierarchy to broader time scales, tones that functioned as 'heads' at lower levels become subservient to other 'heads' at higher time scales. Extending our previous example, the 'head' of the first four chords ultimately is itself subordinated by the final chord of the sequence, as evidenced by the fact that the branch extending from the final chord reaches over the branch from the initial chord. The result is a hierarchical structure of dominance/subservience in tonal relationships as segmented by time-spans, with branches showing elaborations of pitch events. (The well-formedness and preference rules associated with time-span and prolongational reductions are complex, and will not be specified here.)

Prolongational reduction

The function of ***prolongational reduction*** is to express 'the sense of tension and relaxation involved in the ongoing progress of music ... the incessant breathing in and out of music in response to the juxtaposition of pitch and rhythmic factors' (Lerdahl & Jackendoff, 1983, p. 179). Movement away from the tonal center, dissonance, a climactic point of a phrase are all examples of musical tension, while return to the tonic, consonance following dissonance, arrival at the end of a phrase are usually experienced as relaxing. Diagrammed in 'tree' notation appearing above the musical notation, prolongational branching is represented by tensing motions depicted by right branches, and relaxing motion by left branches. The prolongation reduction for Bach's *C Major Prelude* is shown in Figure 7.3.

Unlike the time-span reduction, prolongations (as the name implies) are not constrained by group boundaries. Thus, whereas the resulting time-span units in a time-span reduction are fairly evenly spaced, units in a prolongational reduction can be of varying lengths and can extend across group boundaries. Similar to Schenker's 'prolongation,' a note or chord may remain in effect across measures or even a whole section, even when not physically sounded

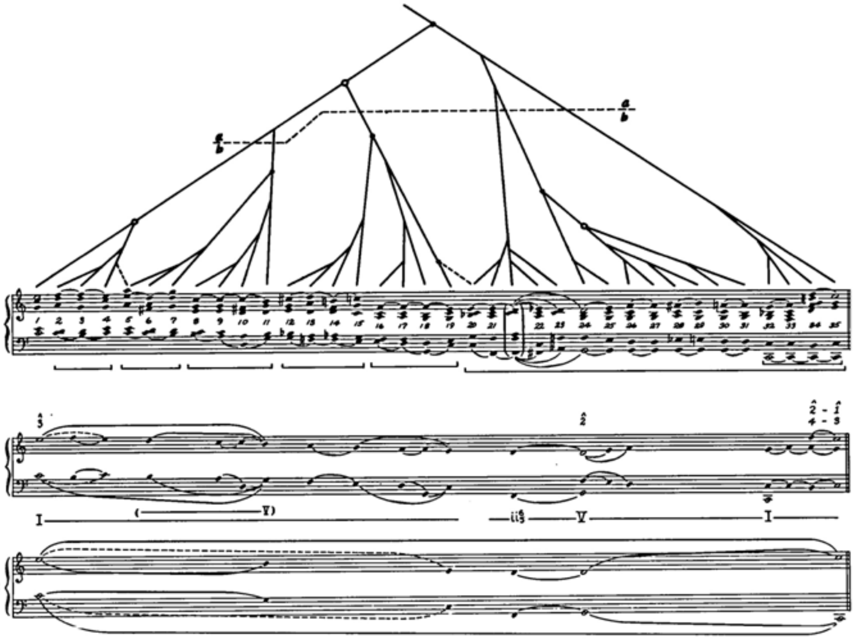


Figure 7.3 Prolongational reduction of Bach's *C Major Prelude*.

Source: Lerdahl, Fred, and Ray S. Jackendoff, *A Generative Theory of Tonal Music*, Figure 10.11, p. 263, © 1983 Massachusetts Institute of Technology, by permission of The MIT Press.

at each moment throughout them. For instance, a tonic triad can be heard as suspending across an entire section, including several contiguous time-spans. This is represented by the 'slur' notation in the reductions shown on the separate staves at the bottom of Figure 7.3. Note that in many respects, the hierarchical structure of time-span and prolongational reductions are similar, as can be seen at higher hierarchical levels in Figures 7.2 and 7.3. However, at lower levels of the hierarchy, these structures diverge; one sees less influence of metrical and grouping structures – hence less influence of time – in prolongational than in time-span reductions.

More recently, Lerdahl (2001) has developed the theory of *Tonal Pitch Space* as a means of quantifying the intuitions behind prolongational reduction. This new theory generates distance metrics among pitches in a tonal system (cf. Krumhansl, 1990), and thus can be used to predict the degree of 'stability' associated with certain pitches (and chords), as well as the degree to which certain pitches 'attract' other pitches to them, as in the way that the tonic note 'attracts' the leading tone to it.

Our brief overview of the theory serves only as an introduction, and has barely captured the complexity of GTTM. In describing the notational system, it is important to remember that it is merely a convenient way of expressing what the theory aims to do – which is to provide a formal description of the musical intuitions of a listener, sorting the 'raw material' at the musical surface of a piece of tonal music into organized 'heard structures.' It is not a theory that focuses on the listener's moment-by-moment response to the tone-by-tone unfolding of a musical work in motion, but rather one that seeks to convey the listener's lasting impressions after a phrase or section or even a whole composition has been presented (i.e., the 'final state' of the listener's understanding; Lerdahl & Jackendoff, 1983, p. 4).

Empirical support for GTTM

GTTM has been tremendously influential on the study of music cognition. The website Google Scholar reports well over 1200 citations of Lerdahl and Jackendoff (1983). In some respects, though, its influence has been felt more as an overall perspective than as a specific set of predictions. In comparison to the myriad citations of the theory, relatively few papers have rigorously tested its principles, although the authors have invited empirical tests of their work. This may be due in part to the fact that the original theory was not designed to frame quantitative predictions (Frankland & Cohen, 2004).

Of the empirical work on GTTM, most has been devoted to its claims concerning preference rules. One of the earliest attempts in this regard was Deliège (1987), who demonstrated that perceived segment boundaries are generally well predicted by GTTM rules for music drawn from the standard repertory. Subsequently, Frankland and Cohen (2004) verified grouping rules with artificially constructed stimuli designed to measure quantitatively the degree to which different principles predict grouping. A major advancement in the quantification of GTTM preference rules is described by Temperley (2001), who instantiated (and extended) these principles in an artificial intelligence system. The theory of Tonal Pitch Space, which evolved out of GTTM's concept of prolongational reduction, has likewise been shown to predict the experience of tension and relaxation in music listening (Lerdahl & Krumhansl, 2007).

The idea that the mental representation of music may be 'reduced' in the sense predicted by GTTM has also seen some support. For instance, influential work by Bharucha (e.g., Bharucha, 1984) has shown that when people hear a tonally unstable note followed by a stable note (see chapter 5), they hear the unstable note as if it were 'anchored' to its more stable neighbor. In other words, the unstable note is treated as somehow incidental. In other research, reductions have been found to predict perceptions of tension and relaxation (e.g., Krumhansl, 1996), closure (e.g., Palmer & Krumhansl, 1987), and memory (Large, Palmer, & Pollack, 1995). At the same time, a problematic result was reported by Cook (1987). He recomposed classical pieces so that they ended in a different key from the key in which they began. According to classic Schenkerian views (consistent with the spirit of GTTM), such alterations should profoundly alter the large-scale structure of the piece. Listeners, however, were apparently insensitive to these large-scale alterations. One possibility is that listeners are only sensitive to large-scale structures at a subconscious level that is hard to measure.

Melodic expectations and the Implication–Realization model

We now turn to a different approach to the cognitive representation of music. Whereas the ultimate goal of GTTM is to elucidate the 'deep' hierarchical structure of music, the next model seeks to explain expectations for events at the musical 'surface.' The emotional experiences that listening to melodies brings about are intimately connected with the way in which a sequence of notes fulfills or violates our expectations (see also chapter 13). Although an earlier theory of how these expectations work was suggested by Leonard Meyer (1956), a more fully formalized and widely tested version was proposed by one of Meyer's students, Eugene Narmour (1990, 1992), in his ***Implication–Realization model***.

The structure of expectations

The central question Narmour asks is: 'What are the specific, note-to-note principles by which listeners perceive, structure, and comprehend the vast world of melody?' (1990, p. 3).

Narmour's Implication–Realization model (hereafter referred to as the I–R model) assumes that a listener's experience of melodic structure is shaped by *expectations of how a melody will continue*. These expectations are influenced by two perceptual systems: one following 'bottom-up' principles that are not dependent on learning and are thought to be universal, and the other applying 'top-down' processes, which are influenced by learning and may be specific to certain broad styles or genres of music, or even more specifically to the works of a particular historical period or to a single composer.

Narmour begins with melodic expectations of the broadest sort: the hypothesis that when presented with two musical events of some sort (e.g., two interval patterns, two pitches in succession, etc.), a listener will subconsciously or consciously infer one of two general outcomes:

(a) *Similarity* (or *process*): From the sameness or similarity in successive musical events, a listener comes to form an expectation of more similarity, which can be expressed in the form of the hypothesis (where → stands for 'implies'):

$$A + A \rightarrow A$$

(b) *Differentiation*: Change in successive musical events leads to the expectation of more change, expressed as follows:

$$A + B \rightarrow C$$

What are the dimensions along which sameness or differentiation may occur? Almost any musical dimension, in fact. According to Narmour, these could be 'pitches and durations,' 'melodic intervals or durational patterns,' and musical forms as 'repeated or differentiated units.'

With respect to more specific melodic implications, on its simplest tone-to-tone level, the I–R model accounts for what we expect the third note to be after hearing two notes before it, assuming that the first two notes do not imply a completion. In this context, the first two notes form such an ***implicative interval***, and the second and third notes form the ***realized interval***, based on the fact that this second interval addresses the implication of the first interval. (Note that in this context, the word 'interval' should be taken to refer to a *melodic* interval rather than a harmonic interval that is simultaneously played.) Realized intervals that match expectations suggest closure and release, whereas realized intervals that deviate from expectations suggest continuation and tension. According to Narmour, 'it is the dynamic, non-reductive, individuated melodic motion on the note-to-note level that captures music lovers' (1992, p. 331).

Given two melodic tones that are not perceived as completed, the listener generates expectations for the third tone, but not every possible pitch is as strongly implied for the continuation of the melody. The expectancies arising from the implicative interval are, in part, influenced by five 'bottom-up' principles of continuation that have roots in Gestalt ideas about grouping. In our discussion of the Gestalt principles of perception in chapter 5, we saw that the Gestalt pioneers were mainly concerned with the visual perception of static figures. However, as Narmour explains: 'in invoking principles of proximity, similarity, and common direction in a temporal art, I hypothesize that rules governing such similarity apply not only retrospectively to perceptual things already seen or heard but also *prospectively* to *expectations* yet to happen' (1990, p. 73, emphases added). Formally, these implicative principles can be set out as follows:

(1) *Intervallic difference*: This principle suggests that after a small interval (five semitones or fewer, a semitone being equivalent to a half-step), the realized interval is expected to be similarly sized; whereas after a large interval (seven or more semitones, or half-steps), the realized interval is expected to switch to a smaller interval. (Intervals of six semitones, which is an augmented fourth or diminished fifth, are considered neither large nor small, and thus lead to neutral expectations.) In other words, intervals are typically expected to be equal to or smaller than a 'perfect fourth' (five semitones), and so when one encounters a large interval ('perfect fifth' or larger), the listener expects the following interval to revert back to the small interval standard.
(2) *Registral direction*: Following a small interval, the listener expects a realized interval that continues in the same pitch direction. That is, an ascending pitch interval that is small generates the expectation of continued ascents to follow. However, after a large interval, the listener expects the next interval to reverse its direction. Thus, a large ascending pitch interval generates the expectation that the next interval will go down.
(3) *Registral return*: A more extreme version of registral direction, registral return suggests that one expects the second interval to bring the pitch back (within two semitones or fewer) to where it began at the beginning of the implicative interval.
(4) *Proximity*: Small realized intervals are favored, because these result in pitches that are close (i.e., proximal) to the end point of the implicative interval.
(5) *Closure*: Narmour emphasizes that this principle does *not* refer to the Gestaltists' idea of 'closure' and need not occur at the end of a phrase or section, but refers more broadly to a weakened sense of implication that varies in degree (Narmour, 1990, p. 11).

Of the principles listed above, the principles of 'registral direction' and 'intervallic difference' play important roles in the I–R model. Moreover, the conjunction of these rules leads to the prediction of a melodic archetype identified by Leonard Meyer as the '*gap-fill principle*' (Rosner & Meyer, 1982). That is, if an implicative interval is smaller than five semitones, a listener expects a continuation of pitch direction. However, if an implicative interval is large, the listener expects a reversal of pitch direction, as well as a smaller interval than was heard implicatively. A good example of the latter situation is the opening of 'Somewhere Over the Rainbow,' which begins with an ascending octave that is directly followed by a one-semitone downwards interval.

Note that in certain cases the principles may conflict with each other. For instance, registral return may predict a large interval at the same time as other principles (proximity, intervallic difference) predict a small interval. According to the model, such clashes are not problematic, but rather contribute to the tension created by a melodic structure. Further, the I–R model does not restrict melodic implication to aspects of pitch and interval alone. For instance, with respect to closure, factors such as tone duration, meter, and harmony may all influence the strength of an implication. These may also determine how weak or strong the expectancy for melodic continuation will be; for instance, Narmour points out that melodic implication can be suppressed by a rest pause, metric stress, or stable tonality.

Bottom-up and top-down expectancies

In Narmour's view, there are two independent 'tracks' operating simultaneously in melodic implication: one which operates on a more automatic and subconscious level (***bottom-up*** processes) and one which interprets incoming data more flexibly within the framework of previously learned knowledge (***top-down*** processes). So far, we have focused mainly on the 'bottom-up' principles, and in particular the application of elementary Gestalt-like principles to tone-to-tone developments of melodic forms. However, 'top-down' differentiations overlay and modify the bottom-up processes described above. These are the influences of one's knowledge and familiarity regarding musical style as a result of exposure or formal training, and consist of *intra-opus knowledge* (of the particular musical work) and *extra-opus knowledge* (of the style of music). Familiarity or formal training in a specific genre of music or a particular historical period or composer serve as sources from which 'top-down' principles shaping an individual listener's expectations may be derived. Each listener possesses a unique body of style knowledge that influences melodic expectations, and this in part may account for variations among the melodic expectancies of listeners.

Ultimately bottom-up and top-down principles jointly determine musical expectations, the former operating subconsciously and on a local level (e.g., an implicative interval), the latter operating consciously and on a more global level (e.g., pre-existing note patterns). Narmour insists on the independence of the two processes as they shape melodic expectancies in separate and distinct ways:

> In terms of *top-down cognition* … a lowly diatonic scale set in quarter notes is a style structure in that, once the listener consciously recognizes its inception, he can easily project its registral, intervallic, metric, and rhythmic continuation and also map out its scale-step conclusion on the basis of his prior knowledge. In terms of *bottom-up perception*, however, the listener also subconsciously, simultaneously, and locally processes each intervallic motion and each registral direction … separately since his initial global projection may be mistaken: The major scale he initially envisions from, say, an ascending C-D-E-F might, for instance, actually continue in a quarter-tone fashion after the F.
>
> (1990, p. 9)

We conclude our discussion with two examples of the application of Narmour's model. Let us first illustrate the application of the basic 'bottom-up' principles, or cognitive primitives of the I–R model, in a simple example shown in Figure 7.4 (from Narmour, 1991, Figure 5, p. 7). The focus here is on the first four notes, or the first three intervals. Intervals 1 and 2 are 'small' according to the I–R theory, and as such one expects a continuation of both registral direction and interval size. However, a surprise comes with the third interval (signified by the exclamation point). Here we have a large interval where a small one is expected. In addition, the large interval continues registral direction, and thus we have *differentiation* on one dimension (interval size) and *process* on another (registral direction). This conflict adds to the surprise and – according to Narmour – enhances the listener's aesthetic experience.

Figure 7.4 'Bottom-up' analysis using the I–R model.

Source: Narmour (1991, Figure 5, p. 7), reprinted with permission from the author and the University of California Press.

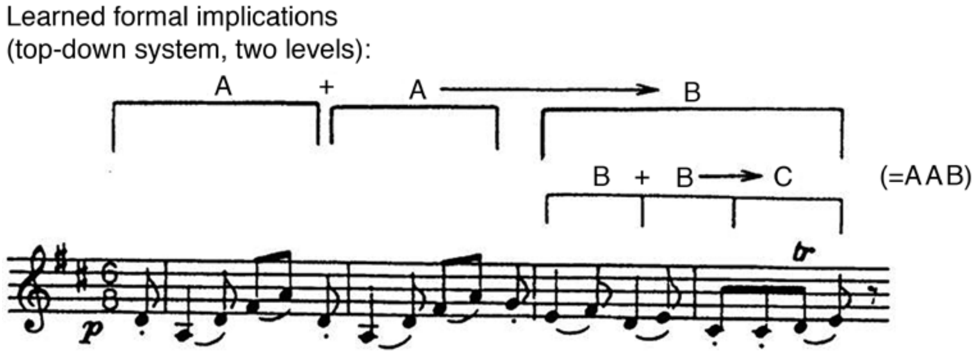


Figure 7.5 The interplay between 'bottom-up' and 'top-down' implications.

Source: Adapted from Narmour (1991, Figure 15a, p. 13), reprinted with permission from the author and the University of California Press.

Through repeated exposure, the listener can form expectations by way of the top-down system that counteract bottom-up tendencies. This leads to a curious prediction of the I–R theory: Music can surprise our top-down expectations while at the same time fulfilling our bottom-up expectations. An example is shown in Figure 7.5. Here we see a motif that is repeated twice; based on *bottom-up* principles, the listener should expect *process* (a repetition of the motif). However, a different motif follows these repetitions. Narmour argues that, over time, a listener comes to expect the differentiation A + A → B within the *top-down* system in a way that counteracts the A + A → A expectation typical of the *bottom-up* system. The example here does just this. In this example, Narmour points to a clever embedding of themes within themes. Within the 'B' segment identified at the top of the graph, there is an embedded theme, B + B → C, that mirrors the A + A → B theme at the higher hierarchical level.

The I–R model is complex and consists of many principles to account for expectations, and the above introduction to the I–R model is necessarily simplified. A full explication of the model requires an eight-page glossary of symbols (Narmour, 1990, pp. 435–442). Narmour's various books and articles have elaborated the model we have sketched in several dimensions. Subsequent research has confirmed and extended Narmour's ideas in significant ways, as discussed in the next section.

Empirical support for the I–R model

Like GTTM, empirical tests of the I–R model have primarily focused on the simpler claims of the model – that is, its 'bottom-up' components. These principles appear to account for melodic expectancies quite well, particularly when combined with a component based on the tonal hierarchy that arises from Krumhansl's (1995) research. For instance, when given two tones and asked to rate how well a third tone serves as a continuation of the melody, the responses were in line with Narmour's predictions (e.g., Cuddy & Lunney, 1995). Similarly, when participants were given two notes and asked to compose the rest of the melody, the third tone they chose in creating their own melodies was well predicted by the I–R model (Thompson, Cuddy, & Plaus, 1997). The level of musical training of participants did not make a difference in these studies, which is consistent with Narmour's assumption that the application of 'bottom-up' principles is not dependent on formal musical training. Further, Krumhansl (1995) and Schellenberg (1996) have shown that these 'bottom-up' expectations can transcend cultural boundaries, and can apply to expectancies for Chinese music and – unlike Krumhansl's tone profile – can also account for expectances in atonal music.

However, in another cross-cultural study, the bottom-up predictions of the I–R model were shown to be culturally constrained (Krumhansl et al., 2000). This research investigated expectancies for Sami *Yoiks*, a style of music originating among the Sami people in Northern Europe that is characterized by large pitch leaps. As one might expect, the I–R model, based largely on an assumption that small intervals are favored, did not provide as strong an account of expectancies for *Yoiks* as it has for other musical styles. Specifically, listeners who were unfamiliar with *Yoiks* made expectancy implications that were better predicted by the model than did listeners who were familiar with *Yoiks*. In other words, Sami listeners had modified the schemata they used to form expectations for *Yoik* melodies.

Narmour's model has also stimulated further theoretical work, refining and expanding models of melodic expectation. Schellenberg (1996, 1997), for instance, demonstrated that the I–R model could be simplified by removing some predictor variables without loss of its predictive power. His parsimonious 'two-factor model' proposes that melodic expectancy is determined by the principle of pitch proximity and pitch reversal, two ideas articulated by Narmour. Drawing on the work of Narmour (1990, 1992) and Lerdahl (2001), Elizabeth Margulis (2005) proposed a model of melodic expectation, which assigns expectancy ratings ('predictions about degrees of expectedness'; p. 664) to melodic events and associates them with listeners' experiences of tension across melodies. Aside from providing a rubric for computing the expectedness of melodic events, the model also 'enables the graphical representation of moment-to-moment expectancy-based fluctuations in affect across a melody's course' (p. 665), capturing the dynamic and real-time aspects of listening to music.

Coda

There are many different ways to take apart a mechanical toy or a clock, a poem, or a musical work. The guiding questions we have in mind, the way we inspect the internal structure, the tools we use, and what we choose to focus our attention on may all lead to different but often complementary discoveries. In this chapter, a few different approaches to understanding musical structure were explored. The Generative Theory of Tonal Music and the Implication–Realization model capture different aspects of the way listeners perceive and comprehend a piece of music, and both have contributed significantly to our understanding of music cognition. However, none of the theorists claims to have set forth a 'complete' or 'finished' theory. Further, the theories do not attempt to 'explain' music, but to capture the complex combination of innate and learned responses in listeners' perception and cognition of a musical work. Drawing from an analogy by psychologist Paul Fraisse, Narmour explains: 'gravity does not explain architecture, but architecture is subject to its law; likewise, perceptual laws do not explain music, but music cannot escape their influence' (1990, p. 4).

The theories described here offer a good starting point toward understanding the perception and cognition of a piece of music. However, many questions ensue: How do musical structures give rise to emotions in listeners? And how does music come to have *meaning* for the listener? These questions are addressed later, in chapters 13 and 14. But first, having explored the perception and cognition of melody, rhythm, and musical structure in the previous and present chapters, we turn our attention to the emergence and development of these and other musical abilities in the next chapters.

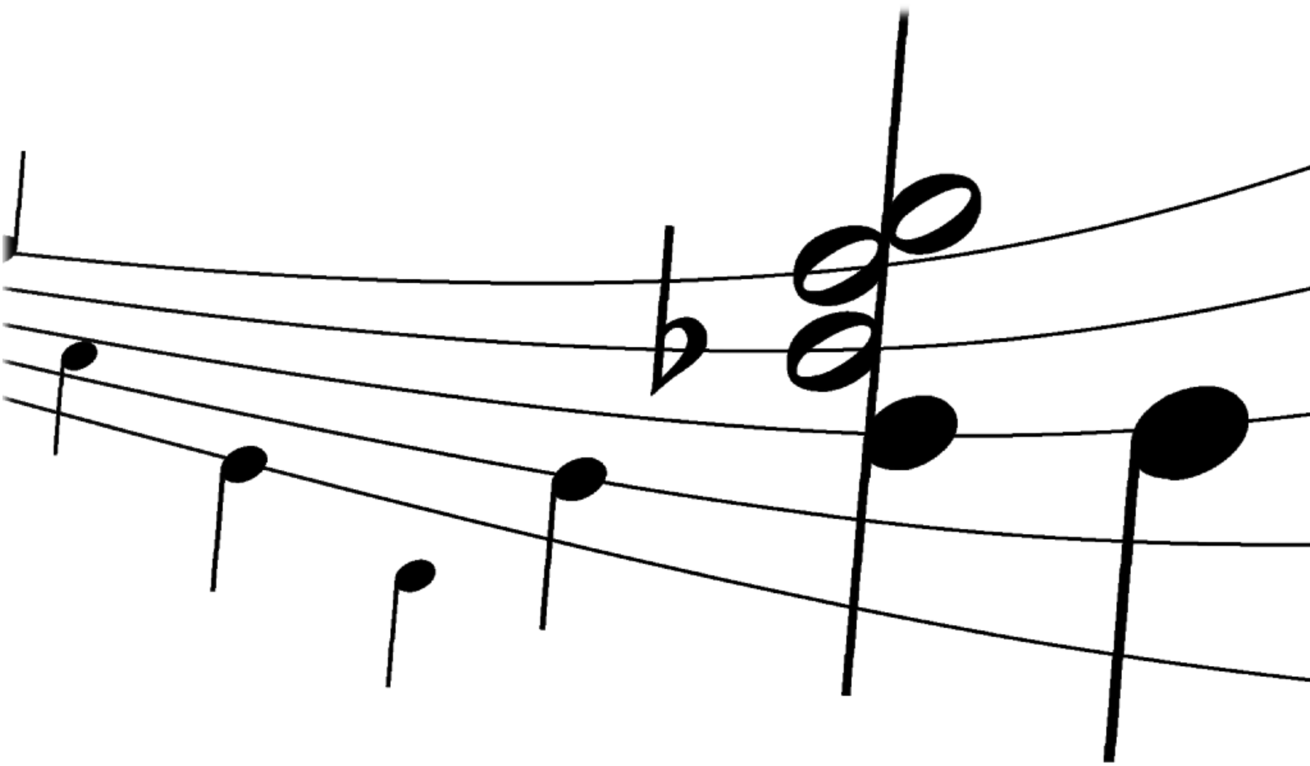
Notes

1 The publication date for Lerdahl and Jackendoff's paper, 'An overview of hierarchical structure in music,' is cited in this chapter as 1983/1984. Note that this is a different publication from their 1983 book entitled *A Generative Theory of Tonal Music*.

2 This rule would lead to the privileged choice of the time signatures 4/4 or common time and 3/4, waltz time.

Part III

Development, learning, and performance



8 Emergence of auditory and music perception

'Let's start at the very beginning / A very good place to start ....' These are the opening lyrics to 'Do-Re-Mi' in Rodgers and Hammerstein's *The Sound of Music*. In previous chapters, we described the auditory system and melody and rhythm perception in the mature adult. In the present chapter, we now retrace our steps to explore the emergence of hearing and music perception in infancy. And we will start at 'the very beginning,' during the prenatal period of development.

This chapter opens with an exploration of sensitivity to sound *before* birth, and presents some intriguing findings in the literature on fetal and newborn responses to auditory stimuli, focusing on the human voice and music. In the second part, our focus shifts to music perception and cognition in infants as we examine some surprising musical capacities they demonstrate in the first year of life. We also address the social significance of music in infants' lives, touching on studies that address the social dimensions of infant music cognition, and caregiver–infant interactions involving expressive speech and singing directed to infants. Throughout, we describe research procedures in some detail, as readers are often curious as to how studies are carried out during the prenatal period or with young infants in the laboratory.

PART ONE: AUDITORY CAPACITIES BEFORE BIRTH AND IN THE NEWBORN

Perhaps no period of human development is as shrouded in mystery as the months between conception and birth. The course of prenatal auditory development is particularly enigmatic, and ideas about what the fetus can hear *in utero* have undergone significant revisions. In the early 1900s, it was commonly assumed that humans had little or no auditory sensitivity before birth. Expectant mothers' reports about movements of the fetus in response to loud environmental sounds were routinely met with great skepticism. Gradually, however, scientific papers were published that began to corroborate these anecdotal reports.

In an early study, Sontag and Wallace (1935) devised a sort of stylus recording device out of inflatable sacks resting against the abdomen, attached to four pens that recorded changes in the shape of the abdomen. When a wood block resting against an expectant woman's stomach was struck loudly, the researchers noted corresponding movements in the contour of the abdomen. Observing this, they deduced that the fetus already responds to external sounds and vibrations before birth. With the advent of technological innovations such as ultrasound technology, much more precise and reliable methods have been devised to study auditory capacities before birth.

Methods for studying fetal responsiveness to sound

Several techniques have been used to present external sounds like music and speech to the fetus. In the ***airborne*** mode, sounds are played via loudspeakers, positioned a few feet away from the mother. In the ***air-coupled*** mode, sound can be transmitted through loudspeakers or headphones placed directly on the mother's abdomen. In the ***vibroacoustic*** mode, a device such as a tuning fork or a voice simulator is placed directly on the maternal abdomen to transmit sounds to the womb.

To examine whether the fetus can detect these sounds, motor responses – such as movements of the head and limbs – can be observed through real-time ultrasound scanning. Physiological responses, such as changes in heart rate following presentation of sound, can also be tracked using fetal heart monitors. A change in heart rate (usually deceleration) is taken as an attentional response, signaling that the fetus detects a sudden change such as the presentation of a sound. Brain imaging techniques such as fMRI (described in chapter 4), but adapted for the fetus (e.g., Draganova et al., 2005; Jardri et al., 2008), have also been used to study prenatal auditory cortical functioning. Brainstem auditory-evoked responses can also be recorded in preterm newborns after birth, in order to detect electrical activity in the cochlea and auditory pathways in the brain, as described in chapter 3.

Studies employing these methods suggest that human auditory functioning begins *about 3 months before birth*. Full-term infants are born at 39 or 40 weeks ***gestational age (GA)***.[1] Most fetuses already respond to sound by around the 26th week GA, as indicated by changes in heart rate or movements in response to loud sound or vibration (e.g., Hepper & Shahidullah, 1994; Joseph, 2000). Brainstem auditory-evoked responses have also been detected in preterm infants at 26 weeks GA or earlier, though becoming more robust by 27 to 28 weeks GA (e.g., Pujol, Laville-Rebillard, & Lenoir, 1998). Within the inner ear, the physical development of the cochlea progresses at a rapid rate during the first two trimesters, and during the third trimester the cochlea becomes 'fine-tuned' for discriminating between specific frequencies (Joseph, 2000; Lecanuet, Granier-Deferre, Jacquet, & DeCasper, 2000), which will be important for the perception of speech and music.

Internal and external sounds in the intrauterine environment

Sounds reaching the womb arise from two kinds of sources. ***Internally generated sounds*** originate from inside the mother's body, such as the mother's breathing, heartbeats, vascular activity, and abdominal gurgles and other digestive sounds (referred to as 'borborygmi' – a good word to remember for your next Scrabble game!). ***Externally generated sounds*** originate from the environment outside the mother's body, such as voices around her, music, ringing phones, and traffic noise. The maternal singing and speaking voice is a special case as it originates internally, but is also transmitted as an airborne sound.

Early recordings obtained with microphones inserted into the uterus yielded internal noise readings of about 72 to 96 decibels (Lecanuet & Granier-Deferre, 1993). Seventy decibels is comparable to the hum of a noisy car engine, and 96 dB is the approximate loudness of a power lawn mower. It was therefore assumed that even if the auditory system were already functional, external sounds such as music and voices would be masked or completely obscured by internal noises. However, early recording methods were not adapted to address fluid impedance. Findings from later studies using ***hydrophones*** (designed to detect sound in underwater environments) range from about 28 to 65 dB (Lecanuet & Granier-Deferre, 1993; Querleu, Renard, Boutteville, & Crépin, 1989, p. 412). Thirty decibels is comparable to the sound level of a quiet library reading room, and 65 dB is the approximate loudness of conversational speech. Thus, the fetus is likely exposed to a richer repertoire of external sounds than previously thought.

Quality of external sounds in the intrauterine environment

External sounds must be fairly loud to reach the fetus, as they must first pass through the maternal abdominal skin and be transmitted through uterine tissue and fluid. This leads to some ***sound attenuation*** (i.e., *loss of intensity* of the auditory stimulus), so the sound reaching the fetus is weaker than the original source.

In general, external speech at a normal conversational level of about 60 dB, and external music presented at a rather loud level of about 80 dB can penetrate through the maternal tissue and emerge above the intrauterine noise level (Querleu, Renard, Versyp, Paris-Delrue, & Crépin, 1988; Gerhardt & Abrams, 2000). However, the quality of sound reaching the fetus also depends on the level of internal noise, which varies with the mother's level of activity. Internal noises such as abdominal gurgling following a large meal, or cardiovascular sounds and movement of amniotic fluid during physical activity, may mask external sounds. If the mother goes to a concert after a hearty dinner, or listens to music while exercising, high internal noise levels may be competing with music from the external environment!

Another variable is frequency, which is related to pitch (see chapters 2 and 3). In general, high-frequency sounds are subject to greater attenuation than low-frequency sounds, as the abdominal wall acts like a ***low-pass filter***. Generally, musical tones up to about 300 Hz (a little above a middle C) pass through to the intrauterine environment virtually unattenuated – that is, with little loss in sound energy (Abrams, Gerhardt, & Peters, 1995). On the other hand, high-frequency sounds, such as tones in the highest octave of a grand piano, may be lost.

Some recordings have been taken from inside the womb to assess the clarity of external sounds reaching the intrauterine environment. For instance, when intrauterine recordings of meaningful and nonsense words were played to adults through loudspeakers at about 60 dB, it was found that only about 30% of the phonetic information was recognizable to adult listeners, although the intonation was well preserved (Querleu et al., 1988). Spoken consonants are higher-frequency sounds than vowels, and are thus subject to greater attenuation than vowel sounds (Gerhardt & Abrams, 2000). Speech therefore loses crispness and articulation by the time it reaches the fetus, sounding rather like a person speaking with a mouthful of marbles!

Few studies of this sort have been conducted using orchestral music. Woodward and colleagues played Bach's *Brandenburg Concerto No. 1* presented at 80 dB, and recorded the sound from inside the uterus (Woodward, Guidozzi, Coley, Anthony, & De Jong, as cited in Woodward, 1992). Although attenuation of the higher frequencies made the music sound very muffled, many of the distinctive qualities of the music – such as rhythm and orchestral texture – were preserved. In another study, an intrauterine recording of a segment of the distinctive first movement of Beethoven's *Fifth Symphony*, opus 67, was easily identified by adult listeners (cited in Abrams et al., 1998).

Fetal responses to music

Fetal responses to music have been investigated in only a handful of studies, but there is some evidence that fetuses respond to music in the last trimester (i.e., last 3 months) before birth. For instance, a piano rendition of Brahms' *Lullaby* evoked changes in heart rate at

28 weeks GA, and increased body movements at 35 weeks GA (Kisilevsky, Hains, Jacquet, Granier-Deferre, & Lecanuet, 2004). Another study found that heart rate deceleration and frequency of movements were more pronounced in 38-week near-term fetuses in response to a Beethoven piano sonata (*The Tempest*) compared to a choral work or rock music (Wilkin, 1996). Of course, the researcher did not conclude that fetuses have a special preference for Beethoven's music, but conjectured that the heightened responses were due to 'dramatic acoustic changes' – such as sharp contrasts in volume, tempo, and texture – that are characteristic of Beethoven's stormy piano sonatas.

The possibility that maternal responses might mediate fetal responses should also be considered. Zimmer and colleagues (1982) played classical and pop music to mothers via headphones placed on the mothers' ears (not abdomens). The researchers found that 34- to 40-week-old fetuses showed an increase in fetal body movements and fetal movements associated with breathing when mothers listened to 25 minutes of either classical or pop music, compared to the control condition in which no music was played. These effects were stronger for the genre of music that the mother reported she preferred.

The authors conjectured that the mother's preference for a piece of music may affect her physiological state, which in turn may affect the fetus' responses. Degree of maternal tension or relaxation has been shown to affect fetal behaviors, including heart rate, respiration, and body movements (van den Bergh, 1992). Further, if the mother's abdominal muscles are relaxed when listening to pleasant auditory stimuli, this may increase the amount of space in which the fetus can move. The mother's role as a possible mediating influence is important to consider in studies in which the mother can hear the music presented to the fetus, especially in light of findings showing that responses to both pleasant and unpleasant music may be enhanced during pregnancy (Fritz et al., 2014).

Links between prenatal exposure and newborn responses

Maternal heartbeat and voice

Given that the fetus shows sensitivity to sound before birth, the question arises whether prenatal exposure to particular sounds can have a lasting influence after birth. For instance, are newborns soothed by the sound of the human heartbeat, as it is a familiar sound from the womb? In a classic study, Salk (1961) found that newborns who heard the sound of a heartbeat at about 72 beats per minute (presented for four days) cried less and were more likely to gain weight during those four days, compared to a control group that were not exposed to heartbeat sounds. This short-term calming effect seems to generalize to other slow rhythmic sounds. Infants have been shown to respond similarly to the sound of the resting heartbeat, lullabies sung in a foreign language, and even metronome clicks at a rate of 72 beats per minute (Brackbill, Adams, Crowell, & Gray, 1966).

Another question often asked is whether newborns are particularly responsive to the sound of their mothers' voices. Indeed, DeCasper and Fifer (1980) found that infants under 3 days old already prefer their own mother's voice to the voice of another female. In this study, newborns sucked on a nipple that was connected to sound equipment through a pressure transducer, so that pressure on the nipple could control the audio track being played. The rhythmic pace of the newborn's sucking produced either a recording of the mother's voice or of another woman's voice. Newborns quickly learned the pattern of sucking that elicited their own mother's voice and produced her voice more often than that of the other female, even though both women were reading the same prose passage.

This familiarity seems to be based mainly on prenatal exposure: Subsequent studies showed that newborns *under 2 hours old* already demonstrate a preference for their own mother's voice over the voices of other females (Querleu et al., 1984). On the other hand, newborns who are 2 days old did not show any preference for their father's voice over the voice of another male (Lee & Kisilevsky, 2014).

In another study, DeCasper and Spence (1986) asked expectant mothers to read one of the following stories out loud every day during the last six weeks of pregnancy: Dr. Seuss' *The Cat in the Hat* (1957), a different children's story, or *The Dog in the Fog* (the last story was created by the researchers by replacing nouns in *The Cat in the Hat* so that the acoustic characteristics of the words were changed, but the number of syllables remained the same). Using the same procedure as in DeCasper and Fifer's study, the researchers found that 2-day-old newborns preferred the story that had been read to them while *in utero*, compared to either of the two stories that they had not been exposed to prenatally. Because the familiar story was preferred even if the recording was made by a female that was not the mother, the researchers concluded that newborns can recognize some acoustic characteristics of speech, such as syllabic stress and voice-onset-time of consonants. Indeed, a subsequent study suggests that by 34 weeks GA, even *fetuses* may already respond to a familiar verse, as indicated by a change in fetal heart rate in response to a rhyme to which they were repeatedly exposed (Krueger, Holditch-Davis, Quint, & DeCasper, 2004).

Vocal and instrumental music

Most studies that have investigated possible links between prenatal exposure and newborn responses to auditory stimuli have focused on responses to speech, and relatively few studies have examined the possible effects of prenatal exposure to music. Using similar methods to the ones devised by DeCasper and colleagues, both Satt (1984) and Panneton (1985) found that newborns preferred a lullaby that mothers had repeatedly sung during the last trimester of pregnancy to an unfamiliar lullaby that the infants had never heard before.

Similar results have been found for fetal responses to external music (e.g., Hepper, 1991; Wilkin, 1996). For instance, Hepper examined a group of 15 mothers who reported that they had watched the show *Neighbours* (a soap opera broadcast on Australian television) every day during their pregnancy. When the *Neighbours* theme song was played to their newborn infants two to four days after they were born, they exhibited a decrease in heart rate, a decrease in movements, and tended to adopt an alert state while exposed to the music. Newborn infants whose mothers had not watched the soap opera before did not exhibit any significant changes in heart rate or movements, and very few became quiet and alert upon hearing the music. However, the effect only lasted for a few weeks, as the newborns' responses to this music did not differ from that of a control group about three weeks later.

Remarkably, infants may even recognize the contour of a melody to which they have been repeatedly exposed before birth (Granier-Deferre, Bassereau, Ribeiro, Jacquet, & DeCasper, 2011). The researchers exposed fetuses to a simple nine-note descending melody twice a day during the last few weeks of gestation, while still in the womb. Six weeks later (after not having heard the melody during the time that ensued), the 1-month-old newborn babies responded with decelerations in heart rate that were twice as large when the same descending melody was played as when they heard an ascending melody with the same tempo, rhythm, timbre, and amplitude envelope (explained in chapter 3). A similar neuroscientific study using event-related potentials (ERPs) also showed that repeated prenatal exposure to a melody for three months before birth induced neural responses to these melodies that lasted for several months after birth (Partanen, Kujala, Tervaniemi, & Huotilainen, 2013).

Beneficial effects of music for the newborn?

It is interesting to learn that infants respond differently to some sounds to which they were regularly exposed during the last few months before birth. However, it must be noted that this apparent 'recognition' of familiar materials is short term, lasting only a few days to several months after birth (e.g., Fifer & Moon, 1989). As yet, there is no hard evidence that prenatal exposure to music or other auditory stimuli 'teaches' the fetus something that will carry through to childhood and beyond, or primes them for exceptional ability.

Similarly, the idea that there are lasting benefits of playing classical music to newborns and infants – and in particular, that doing so can boost cognitive functioning – has not been demonstrated in the scientific research. This popular notion, and the way it came to be associated with the '***Mozart effect***' (the enhancement of cognitive skills through listening to classical music, particularly the music of Mozart), has an interesting history.

The origin of the so-called 'Mozart effect' can be traced back to a brief paper by Frances Rauscher, Gordon Shaw, and Katherine Ky which appeared in *Nature* in 1993 (followed by a longer 1995 article by the same authors). The study showed that listening to Mozart's *Sonata for Two Pianos in D Major* (K. 448) for 10 minutes enhanced adult listeners' performance on standard spatial tasks, compared to listening to a relaxation tape or to nothing at all. Rauscher and colleagues found that those who had listened to the Mozart sonata immediately before completing one of three spatiotemporal reasoning tests taken from the Stanford–Binet intelligence scale performed significantly better than those in the other conditions. However, the effect was short-term, 'washing out' after about 10 to 15 minutes.

Rauscher and her colleagues were cautious in interpreting the findings of their study, and took pains to point out its limitations. However, when findings of the study were disseminated in the media – where they were dubbed as the 'Mozart effect' (a phrase the researchers did not use in their original paper) – the media reports often seemed to imply or even directly claim that listening to Mozart can permanently boost intelligence or IQ! The facts that the original study focused on a specific type of intelligence (spatial reasoning), that the effect lasts only about 10 to 15 minutes, and that subsequent studies have usually failed to replicate the original findings (e.g., Chabris, 1999) were often omitted.

Other explanations for the original findings have received far less attention in the media. For instance, the original interpretation of the Mozart effect was that listening to Mozart stimulates neural networks that are also used in spatial reasoning. However, Nantais and Schellenberg (1999) pointed out that the three conditions originally used (Mozart, silence, relaxation tape) did not merely differ with respect to their ability to stimulate this network, but with respect to the boredom they caused in the listeners, which was substantially less for Mozart's music than the other conditions. Rather than stimulating a specific neural network, subsequent studies suggest that the temporary enhancement of performance associated with Mozart's music may be due to its effects on the *arousal or mood* of listeners (Thompson, Schellenberg, & Husain, 2001), and this effect may therefore be brought about by a range of other stimuli, including music by other composers, music of non-classical genres, and even recorded prose passages (for a review, see Schellenberg, 2012).

One of the unexpected outcomes of this research is that the Mozart effect somehow became linked in public perception with the idea of wide-reaching benefits of classical music *for infants*. This shift can be traced in news media reports linked to this study. Social psychologists Bangerter and Heath (2004) analyzed 478 articles related to the Mozart effect appearing in the top 50 American newspapers from the day the original *Nature* article was published in 1993 until 2002. They found that the proportion of newspaper articles related to the topic of the Mozart effect that mentioned infants increased from 0% in 1995 (the year the longer report was published) to about 55% of news reports by 1999, even though the original studies did not include any infants. Meanwhile, college students – the participants in the original study – were mentioned in 80% of articles related to the Mozart effect in 1994, and this steadily decreased to only about 30% after 2000!

Articles in the news and popular media with titles such as 'Mozart makes you smarter, California researchers suggest' (Knox, 1993) and 'More proof music lifts young IQs' (Ingram, 1997) reinforced the idea that scientists had established a link between music and enhanced intelligence in the developing brain. In 1998, the Governor of the US state of Georgia even proposed a budget to supply classical CDs to all parents of the approximately 100,000 infants born each year in that state (Sack, 1998).

Today, the notion of the so-called 'Mozart effect' continues to be used for marketing purposes. Books, audio recordings, videos, and other products claiming that classical music can 'make babies smarter' are still available, though not as pervasive as they once were. To date, however, there is no consistent body of scientific evidence that has demonstrated a strong or direct connection between short-term exposure to classical music and enhanced cognitive functioning in infants, and many attempts to replicate the Mozart effect in children have yielded no effect (e.g., Črnčec, Wilson, & Prior, 2006; McKelvie & Low, 2002; cf. Ivanov & Geake, 2003). Only prolonged active engagement in music, such as learning to play a musical instrument, has been associated with long-term effects on general intelligence in some studies (e.g., Rauscher & Hinton, 2006; Schellenberg, 2006; though see Costa-Giomi, 2004). Thus, the Mozart effect can be viewed as 'a scientific legend' that drew credibility through association with a scientific study, but eventually 'transformed to deviate in essential ways from the understanding of scientists' (Bangerter & Heath, 2004, p. 608).

That is not to say, however, that music may not have other kinds of positive effects on young infants. Research findings such as the ones reviewed in this chapter have been applied to some innovative interventions with at-risk infants in neonatal intensive care units (NICUs). For example, benefits have been found for interventions employing 'biological maternal sounds' (BMS), which consist of recordings of the maternal heartbeat at rest and the mother's voice while both speaking and singing songs to her infant. These sounds are treated with a low-pass filter with a cut-off at 500 Hz, to simulate how the maternal heartbeat and singing and speaking voice might sound from inside the womb. A few short doses of BMS per day – played through an MP3 player inside an infant's isolette or crib – have been shown to improve the regularity of breathing (Doheny, Hurwitz, Insoft, Ringer, & Lahav, 2012), facilitate faster weight gain (Zimmerman, Keunen, Norton, & Lahav, 2013), elicit a soothing response (Rand & Lahav, 2014), and stimulate development in the auditory cortex (Webb, Heller, Benson, & Lahav, 2015) in preterm and very low birth-weight infants, compared to exposure to other recorded environmental sounds.

A primary area of concern pertains to the noise levels in NICUs, especially with regard to high-frequency noise. Whereas fetuses are exposed to mainly low-frequency sounds *in utero*, the NICU is filled with many high-frequency sounds coming from ventilators, monitors, and alarms. Newborns are subjected to frequencies in the 500 to 3150 Hz range (around B4 to G7 on a grand piano) for prolonged periods (Lahav, 2014) at loudness levels of 60 to 80 dB. Therefore, one practical function of music is that it can provide a continuous and relatively low-frequency sound that can mask other noises, thus reducing stress and avoiding disruptions of sleep. Indeed, studies have shown that soft live singing or soft recorded instrumental music in NICUs improves the quality of sleep, decreases stress-related behaviors, alleviates pain after procedures, reduces weight loss, and positively affects cardiac and respiratory functions (e.g., Cassidy & Standley, 1995; Loewy, Stewart, Dassler, Telsey, & Homel, 2013; Tramo et al., 2011; cf. Cevasco, 2008). Whereas there is no clear evidence that early exposure to music 'boosts brainpower' in infants, music may offer a range of other important benefits.

Many therapeutic uses of music for at-risk infants are explored in practical guides such as *Music Therapy for Premature and Newborn Infants* (Nöcker-Ribaupierre, 2004), *Music Therapy with Premature Infants* (Standley & Walworth, 2010), and *Music Therapy and Parent–Infant Bonding* (Edwards, 2011). At the same time, further research is needed to test the efficacy of prenatal and postnatal musical interventions before their use becomes widespread (see Adachi & Trehub, 2012).

PART TWO: MUSIC PERCEPTION AND COGNITION IN THE INFANT

William James (1890, p. 488) used the phrase 'one great blooming, buzzing confusion' to describe the world as it might be experienced by the newborn. Like many others in his day, he assumed that infants were poorly equipped to process the rich information flooding their senses. This view has changed dramatically with accumulating evidence of infants' acute sensory capacities. The first part of this chapter outlined basic auditory responses before birth and in the newborn. In the second part, we will focus on music perception and cognition in the infant from birth to about one year of age. By the end of our discussion, we should have greater insight into why psychologists have replaced the old Jamesian view with the idea of the 'sophisticated infant,' already handily equipped to process complex sensory stimuli.

Orientation to sound

The newborn already responds actively and curiously to sound, as seen in the tendency to turn toward the source of a sound shortly after birth. Spontaneous head-turns towards the sound of a rattle presented to the left or right side have been observed in 2- to 4-day-old newborns (Muir & Field, 1979). By 6 months, most infants can turn their heads within 4–6° of a sound source (Morrongiello & Rocca, 1987) and by 7 months, infants can even reach for sound-emitting objects in the dark, relying solely on their sense of hearing in the absence of visual cues (Clifton, Perris, & Bullinger, 1991).

Infants' ability to localize sound does not match adult performance, however, when it comes to identifying the sources of more complex sounds in the real world. As discussed in our discussion of ***localization*** in chapter 3, our ability to infer where a sound is coming from relies on many cues. Two of the binaural cues are ***interaural timing difference*** (if a sound reaches one ear before the other, the source is judged to be closer to that ear) and ***interaural intensity difference*** (if the sound is louder in one ear than the other, it must be closer to the source).

Infants' poorer performance on sound localization tasks may be due in part to the small size of infants' heads. Figure 8.1 shows the head circumference of a newborn infant, taken a

Figure 8.1 First measurement of a newborn's head at birth (~13.5 inches, taken directly after birth). The small size of infants' heads makes it difficult to localize where sounds are coming from based on interaural cues.

Source: Photograph used with permission from the parents.

few minutes after birth. Though disproportionately large for the body (newborns' heads are about 1:4 to body mass whereas the adult ratio is closer to 1:8), the distance between the two ears is much smaller than in adults. Thus, interaural differences in timing and intensity are more difficult to detect. Infants' heads may also not be large enough to produce the 'sound shadow' that conveys cues as to the location of sounds due to sounds having to travel around the head and being absorbed by the head. Further, as young infants spend a lot of time lying on their backs, this too would affect accuracy in detecting sources of sound, as sound localization is more difficult when supine than when standing.

Perception of pitch and melody

Pitch perception

The fetus is mainly exposed to low-frequency sounds while *in utero*, but after birth, infants are more attentive to higher-pitched sounds. By 6 months, most infants can discriminate between two high-frequency tones at similar levels of accuracy to adults (Olsho, Koch, & Halpin, 1987). Perception of low-frequency sounds, however, advances much more slowly, and only approaches adult performance by about 10 years of age. Improvement in hearing acuity can be generally attributed to the growth of the external ear and the increasing efficiency of middle ear conductance from infancy through adulthood (Keefe, Bulen, Campbell, & Burns, 1994).

One line of inquiry concerns how the brain processes simple pitch differences. Although there is evidence that this perceptual ability seems to be present early based on what we know from behavioral data, it is possible that the brain's response evolves over time. In fact, this does seem to be the case, based on measurements of event-related potentials (see chapter 4) in infants. Researchers at McMaster University led by Laurel Trainor have measured EEG responses in infants by placing electrodes at strategic locations on the scalp, to monitor and record brain activity directly under each point of contact. A series of studies by this group measured ERP responses to pitch in 2- and 4-month-old infants, and in adults (He, Hotson, & Trainor, 2007, 2009). In these studies, every trial constitutes the presentation of just a single tone, and usually one hears the same tone again and again. Every so often, the pitch changes unexpectedly. At issue was how the brain would respond to the 'mismatch' of these occasional pitch changes, referred to as 'oddball' tones, to the prevailing monotone context. It was found that the brains of 4-month-old babies responded very rapidly to oddball tones, similar to the responses of adults. However, 2-month-olds did not exhibit the same rapid responses, even though their brains still responded to the mismatch. Thus, neural responses to unexpected tones speed up as our expectations become more fine-tuned.

Consonance and dissonance

Some studies suggest that sensitivity to consonance and dissonance emerges at a very early age. Two-month-old infants prefer ***harmonic intervals*** (i.e., pitch intervals formed by tones played simultaneously) that are consonant, as opposed to those that are dissonant (Trainor, Tsang, & Cheung, 2002). They spend more time looking in the direction of a loudspeaker and listening when it is playing intervals such as perfect fifths (consonant) versus minor ninths (dissonant). Even 2-day-old newborns prefer consonant Mozart minuets to a version with many dissonant intervals, created by altering two notes in the composition consistently by just one semitone (Masataka, 2006). These findings are consistent with claims that the consonance of simultaneous tones likely reflects the mechanics of the cochlea (Plomp & Levelt, 1965, as discussed in chapter 3). Some studies also point to an early sensitivity to consonance in ***melodic intervals*** formed by *successive* rather than simultaneous tones (such as playing a C tone followed by a G tone, versus a dissonant melodic interval such as C followed by an F#) by about 6 months of age (Schellenberg & Trehub, 1996).

However, recent findings are raising questions about theories of consonance (e.g., McLachlan, Marco, Light, & Wilson, 2013, as discussed in chapter 3) and about the assumption of innate preference for consonance (Plantinga & Trehub, 2014). For instance, although some studies seem to point to an early preference for consonance over dissonance, this preference is not apparent in children until about 9 years of age (Valentine, 1962; except in a population with more musically trained children).

In a series of studies, Plantinga and Trehub (2014) found that, in contrast to previous findings, 6-month-olds did not show a preference for consonance over dissonance in unaccompanied consonant versus dissonant melodies, nor melodies accompanied by mostly consonant or mostly dissonant chords. Further, they conducted additional studies employing the same musical stimuli used in previous studies, and found no evidence of preference for consonant over dissonant stimuli in 6-month-old infants. Most interestingly, they found that exposure to only three minutes of either a consonant or dissonant melody with corresponding consonant or dissonant chord accompaniment led to subsequent preference for whatever melody had been played – regardless of whether it was consonant or dissonant. Therefore, even brief exposure to consonant or dissonant music was sufficient to shape the infants' preference.

Thus, evidence for innate or early-emerging preference for consonance may not be as strong or consistent as previously assumed. Future studies may shed more light on the nature of dissonance and the question of whether preference for consonance is innate. We may also gain deeper insight into what ethnomusicologists have long observed: that dissonance is not universally judged to be unpleasant, but is a facet of music that is regarded neutrally or even valued in many musical cultures.

Sensitivity to scales

What about pitch perception in the context of a particular scale system? Studies show that Western adults are much better at detecting tuning changes in melodies written in familiar scales (e.g., major or minor) versus unfamiliar scales. Young Western infants, on the other hand, discern mistuning equally well in melodies in either scale (Lynch, Eilers, Oller, & Urbano, 1990) until about 12 months of age, after which they show greater sensitivity to tuning changes in melodies based on a major scale, as do Western adults (Lynch & Eilers, 1992). Studies suggest that more complex knowledge of key membership and tonality becomes stable by 4 or 5 years of age (e.g., Trainor & Trehub, 1994), though these abilities may emerge in children as young as 3 years, and are facilitated by music lessons even at the beginning level (Corrigall & Trainor, 2009).

A particularly useful feature of scales, which can be found across cultures (discussed further in chapter 15), is the fact that most scales used to create melody are asymmetric with respect to the spacing of chromas (see chapter 5). Trehub, Schellenberg, and Kamenetsky (1999) tested how well infants and adults could learn melodies based on three different scales: the diatonic major scale (most common in Western music), a novel scale created by the experimenters that consisted of asymmetric intervals (similar to the diatonic scale), and another made-up scale with equivalent 'whole-tone' intervals. Whereas adults showed an advantage for the Western diatonic scale, infants performed equally well with either asymmetric scale. Neither group performed well with the whole-tone scale, which has equal spacing between each tone.

Perception of melody

Of course, infants do not simply respond to single pitches in isolation. They follow the general shape of rising and falling tones that forms a melody.

In an important early study, Hsing-Wu Chang and Sandra Trehub (1977) wanted to know whether 5- and 6-month-old infants can recognize a melody if it is played in a new key. Or, as every tone of the melody would be changed (for instance E-G-C in C major would be B-D-G in G major), would the infants simply think it is a 'new' melody? The researchers employed the ***habituation procedure***, commonly used in infant studies. In this procedure, an auditory stimulus is repeatedly presented and then switched to another stimulus to see if the infant detects the difference. The first step is to take the infant's heart rate to establish a baseline measure. Then an auditory stimulus is repeatedly presented. An infant's heart rate tends to *decrease* (signaling interest or attention) when first attending to a new sound, but gradually returns to the baseline rate when the infant gets used to it (i.e., ***habituates***) after it is repeated many times. After the heart rate returns to baseline, the recording suddenly switches to a similar but new auditory stimulus, which is then repeatedly presented. If the infant's heart rate *decreases after the sound is switched* – departing again from the baseline level – the response would suggest that the infant detects that it is a new or different sound from the one before. No change in heart rate would suggest that the infant does not discriminate between the two stimuli.

In Chang and Trehub's study, the same six-note melody was presented repeatedly to infants. Then, either an exact ***transposition*** of the original melody (the entire melody shifted up or down by three semitones) or a 'scrambled' melody (the same six tones of the original melody, but presented in a random order) was played. They found that the infants' heart rate did not decrease when the transposed melody was played. In other words, infants responded to the transposed melody as if it was basically the same melody as before (just as we would hear 'Happy Birthday' in the key of F or G major as the 'same' tune, even though every note of the melody changes when it shifts key). However, heart rate did decrease when the scrambled melody was played, indicating that the infants detected that it was different from the original melody. Just like adults, infants responded to scrambled melodies as if they were 'new' or 'different' from the initial melody.

Trehub and colleagues concluded from further research that infants may not attend primarily to the specific pitches of a melody (e.g., A, D, F, etc.), nor to the exact intervals between the pitches, but to the general ***melodic contour*** (i.e., the overall pattern of rising and falling pitches that make up the shape of a tune, even if the exact size of intervals are changed). This may be why most infants can detect the difference between six-note melodies that differ only in one single pitch if it changes the melodic contour (Trehub, Thorpe, & Morrongiello, 1985), but often cannot discriminate between two melodies that differ in as many as three or more tones if the overall melodic contour is preserved (Trehub, Bull, & Thorpe, 1984).

The research cited thus far might seem to suggest that infants pay attention to relative pitch rather than absolute pitch. However, subsequent research suggests that infants use *both* relative and absolute pitch information, depending on the conditions (Volkova, Trehub, & Schellenberg, 2006; Plantinga & Trainor, 2005). For instance, infants seem to rely on relative pitch in tasks employing short melodies presented in musical context. However, memory for absolute pitch seems to prevail when infants have to learn a sequence of pitches that may not conform to the musical rules usually employed in composition (Saffran & Griepentrog, 2001; Saffran, 2003). In comparison, most adults rely on relative pitch information for both of these kinds of tasks.

Perception of rhythm and meter

Just as infants are able to recognize a melody when it is played in another key, they also recognize rhythms played at different speeds. Sandra Trehub and Leigh Thorpe (1989) found that 7- to 9-month-old infants can differentiate between simple groupings of sounds or rhythmic patterns (e.g., XX_XX versus XXX_X) if the *tempo* (speed) at which they are played is changed, as long as the relative durations between each tone are preserved. Even when the pitches are changed, infants can differentiate between rhythmic patterns that are similar, but not the same. The tendency to focus on *relative* pitches between tones in a melody and *relative* durations in rhythms (as opposed to absolute pitches and absolute time durations) allows listeners to recognize a song even if transposed to a different key or played at a different tempo.

Most music not only has an underlying beat or pulse that can be extracted by the listener, but also a regular pattern of strong (accented) versus weak (unaccented) beats. Even 2- and 3-day-old newborns expect a clear marker at the beginning of a repeating rhythmic cycle (i.e., a 'downbeat'). This was shown in a study using electroencephalography (EEG) to monitor electrical activity in the brain, by placing electrodes on selected points on the infants' heads. When the researchers played a simple rock rhythm with missing downbeats to newborns and monitored their electrical brain signals, the newborns' neural responses were consistent with the brain activity that occurs when someone registers a violation of sensory expectations (Winkler, Háden, Ladinig, Sziller, & Honing, 2009).

Movement appears to play an important role in interpreting the meter of music, as demonstrated in an innovative study by Jessica Phillips-Silver and Laurel Trainor (2005), illustrated in Figure 8.2. In this study, a recording of two minutes of a rhythm pattern that had no accented beats was played to 7-month-old infants. As the music played, researchers held the infants in their arms and bounced them either to every second beat or every third beat of the rhythm. Therefore, although all infants heard the same unaccented rhythm, the 'felt'

Figure 8.2 Conditions used in Phillips-Silver and Trainor's (2005) study: (a) bouncing while watching experimenter; (b) bouncing while blindfolded; (c) passively observing experimenter; (d) preference test.

Source: Copyright © Jessica Phillips-Silver, Laurel Trainor, and McMaster University.

accents were different. During the test phase, two versions of the same rhythm pattern the infants had heard before were presented – but this time, the rhythm was clearly accented either on every second or every third beat (i.e., duple or triple meter). Phillips-Silver and Trainor found that infants who had been bounced in duple time preferred the version that was accented on every second beat, as indicated by longer attentive listening time. Infants who were bounced in triple time preferred the version that was accented on every third beat.

Blindfolding the infants while they were being bounced to the ambiguous rhythmic pattern produced the same results. Again, the infants showed a preference for the rhythm presented with an accent pattern matching the meter to which they had been bounced before. However, when babies *merely observed* an adult jumping in duple or triple pattern to the ambiguous rhythm while the infant sat still, the infants showed no clear preference for the duple or triple pattern audio track. Thus, visual cues alone do not seem to play an important role in infants' representation of beat structure. The infant has to feel the beat with the *movement of his or her own body*, thus incorporating both auditory and vestibular cues in interpreting rhythmic structure. Similar studies carried out with adult participants have yielded parallel findings, showing that they too 'hear what the body feels.' Phillips-Silver and Trainor point to a cross-modal interaction between body movement and auditory interpretation of rhythmic structure that emerges early in infancy and is maintained in adulthood: '*how we move* will influence *what we hear*' (2007, p. 535, original emphases).

Another study by Gerson and colleagues also demonstrated the role of active experience in shaping perception (Gerson, Schiavio, Timmers, & Hunnius, 2015). In two experiments, the researchers found that 6-month-old infants who had *actively played* with a small drum before watching a video gazed at the screen longer if it showed someone drumming synchronously rather than asynchronously with an audio track of the drumming sounds. Only five minutes of active engagement with the drum before watching the video seemed to increase the infants' sensitivity to audiovisual synchrony when later observing someone else playing the instrument. In contrast, infants who had only *observed* someone playing the drum for five minutes prior to watching the video did not show this preference (the two conditions are shown in Figure 8.3). Just as Phillips-Silver and Trainor (2005) had shown, active experience provides a benefit beyond observation, but this study extends the findings to the implications for the perception of synchrony of multisensory stimuli. The authors propose that 'sensorimotor activity consolidates the sensitivity to audiovisual synchrony, promoting a more *embodied* view of the processes involved in learning' (Gerson et al., 2015, original emphasis).

Implications for social synchrony

Intriguingly, subsequent studies have found that interpersonal motor synchrony that is coordinated by music facilitates prosocial behaviors, such as helping and cooperation. Fourteen-month-old infants were bounced to the beat of music by an assistant, while facing an experimenter bouncing either in synchrony or out of synchrony. Infants who had been bounced in synchrony with the experimenter were subsequently more likely to help her when she pretended to drop some items than infants who had watched the experimenter bouncing 'out of sync' with their own movements (Cirelli, Einarson, & Trainor, 2014). Infants were also more likely to help a new unfamiliar adult they perceived to be a 'friend' of the experimenter (after previously watching them interact and solve a problem together), but not a stranger who was not affiliated with the experimenter (Cirelli, Wan, & Trainor, 2016).

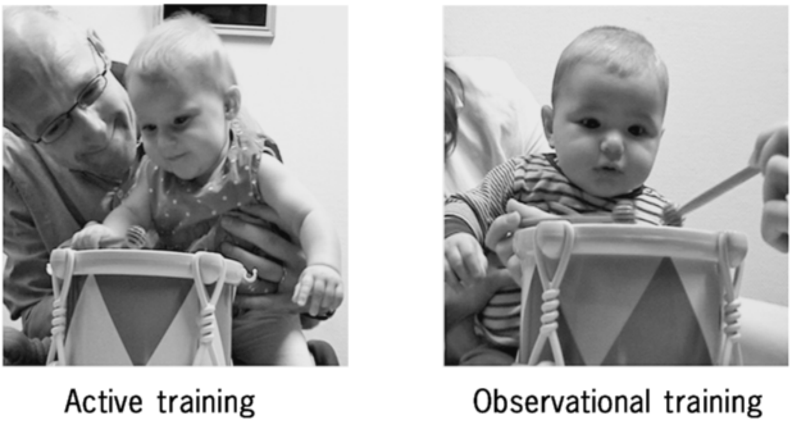


Figure 8.3 The training conditions in Gerson and colleagues' (2015) studies. Infants in the 'active' condition played with the drum, whereas infants in the 'observational' condition observed drumming for five minutes prior to watching videos of a drummer.

Source: Creative Commons Attribution License © 2015 Gerson, Schiavio, Timmers, and Hunnius.

These findings dovetail with a larger body of research on interpersonal synchrony, showing that both adults and children are more likely to like, trust, and cooperate with each other after engaging in synchronous than non-synchronous action (e.g., Hove & Risen, 2009; Kirschner & Tomasello, 2010; Wiltermuth & Heath, 2009), even if the task is as simple as tapping one finger in time to the same beat. Accordingly, singing and dancing together are typical elements of events that are intended to strengthen affiliation between people – including patriotic and religious gatherings, and parties.

The role of experience

Sensitivity to particular meters seems to be shaped by experience, as shown in a study by Hannon and Trehub (2005a) that was briefly mentioned in chapter 6. Hannon and Trehub found that when Balkan folk melodies of Serbia and Bulgaria were played to North American adults, the adults could differentiate between changes in the melody that deviated from the original metrical structure versus those that preserved the meter. However, they could only do so for Balkan melodies presented in simple meters (which are prevalent in Western music) and not for those in complex meters (which are rare in Western music). North American infants of 6 months, on the other hand, were able to discriminate between meter-disrupting and meter-preserving deviations from the original Balkan folk melody in both simple and complex meters.

By 12 months old, however, Western infants performed similarly to Western adults. That is, 12-month-old Western infants were able to tell the difference between melodies that were slightly changed, but only for melodies presented in simple meters common in Western repertoire (Hannon & Trehub, 2005b), and no longer in the complex meters typical of Balkan music. Further, the researchers found that after only two weeks of listening to CDs of folk music with complex meters at home, the 12-month-old infants performed the task well for music in both simple and complex meters, whereas adults showed no improvement on complex meters. It seems that infants start with a flexible style of processing tonal and metrical structure, which takes on a more culture-specific character especially after about a year of maturation and exposure to a particular musical repertoire. However, even at 1 year of age, infants remain malleable and are able to quickly adjust to other metrical structures after only limited exposure.

More basic inclinations toward certain meters may be apparent at an earlier age. Even 4-month-old infants already show preference for meters that are common in their own musical cultures as opposed to unfamiliar meters (Soley & Hannon, 2010; see also chapter 15 in this volume), as demonstrated by the time they spend attending to each when given both options. Thus, preferences or attention to certain metrical structures may emerge quite early in development, before culture-specific biases are manifested in discrimination tasks that reflect perception and processing of meter in ways that may be shaped by one's experience with a particular musical tradition.

Memory for music

Thus far, the studies we have discussed have examined infants' immediate responses to musical stimuli shortly after they were presented. But do infants retain music after a period of time has passed? Some years ago, one of your authors (ST) met an 8-month-old infant whose parents claimed that he would usually stop crying at the sound of the chorus of the pop song 'Y.M.C.A.' Whenever the recording was played while he was distressed, the infant would cry only intermittently, or cease crying altogether. Although his parents tried to soothe the infant with other music, the infant appeared to have a preference for 'Y.M.C.A.' by the Village People. Can 8-month-olds really recognize particular songs?

Addressing the question of young infants' memory for music under much more controlled conditions, Jenny Saffran and colleagues asked parents to play the slow movements of two piano sonatas by Mozart to 7-month-old infants once a day for 14 days (Saffran, Loman, & Robertson, 2000). After daily exposure to the music for two weeks, parents were instructed not to play the recordings for two weeks. The infants were then brought to the lab for testing.

Each infant was tested individually, using an infant research method referred to as the ***head-turn preference procedure***, commonly used in infant studies on auditory perception. In this procedure, a music recording is played when the infant turns his or her head to one side, and switches to another music track when the infant turns to the other – much like a 'juke-box' controlled by the infant's head movements. The total amount of time that the infant spends listening attentively to each piece of music is measured, and taken as an indicator of the infant's interest in the music. The researchers used two 20-second excerpts of the Mozart sonatas to which they had been exposed ('familiar' music) and two 20-second excerpts of new excerpts taken from the same CD recording (the slow movements of two other sonatas by Mozart) for the 'novel' music.

The researchers found that the infants listened significantly longer to the 'novel' Mozart piano sonatas than to the 'familiar' Mozart piano sonatas. Earlier in the chapter, we discussed how newborns are often calmed or soothed by familiar stimuli. However, preferential listening tasks gauge older infants' attention and interest, which is usually oriented to stimuli that are *novel*. We can infer that the infants in Saffran and colleagues' (2000) study recognized the 'familiar' music, as they showed significantly less interest in the music that they had been repeatedly exposed to two weeks earlier than in the 'new' pieces of music. A control group of infants who had never heard any of the Mozart piano sonatas did not show a significant preference for any of the excerpts.

The CD recordings that the researchers gave to the infants' parents were filled with 10 minutes of Mozart piano sonatas. Thus, the infants seemed to have retained fairly long and rich

Figure 8.4 Saffran and colleagues (2000) found that 7-month-old infants retain pieces of music in memory following daily exposure. Is it possible that Zev, shown here at 7 months, recalls some of the pieces his sister plays during her daily practice?

Source: Photograph used with permission from the parents of the children. Copyright © Jessica Phillips-Silver.

pieces of music in their memory following a two-week delay. Further, the infants could not rely on acoustic differences in the music (such as timbre) because the familiar and novel excerpts were all taken from the same CD of Mozart sonatas, all performed on the piano. Subsequent studies have also demonstrated infants' ability to retain music in long-term memory after a delay (Ilari & Polka, 2006; Mehr, Song, & Spelke, 2016; Trainor, Wu, & Tsang, 2004), even when employing complex music such as piano pieces by Ravel. Thus, the research on infant memory for music lends some credence to the parents' claim in our 'Y.M.C.A.' anecdote!

A social dimension to memory

In an intriguing paper entitled 'For five-month-olds, melodies are social,' Samuel Mehr and colleagues explored a possible social dimension to infant memory for song (Mehr et al., 2016). The researchers used two short songs with exactly the same lyrics and rhythm, differing only in the melody. The song was presented in one of three ways: (1) some infants heard the song sung frequently at home by a parent; (2) other infants heard the song played through an audio recording embedded in a stuffed animal that the parent brought home; and (3) others heard the song sung by a friendly adult via interactive video sessions at home. All infants heard the song frequently in the assigned fashion in their own home for one to two weeks.

When the infants returned to the lab, they were shown a video of an unfamiliar adult singing one song, and another unfamiliar adult singing the other song (same lyrics and rhythm, but different melody). Mehr and colleagues found that the infants looked significantly longer at the new adult who sang the song they had heard repeatedly at home, *but only if the song had been sung to them by their parent*. Infants who had heard the song played by the stuffed animal at home or through video sessions did not show a preference for the new individual singing the familiar song. The researchers speculated that the infants' heightened interest in an unfamiliar adult singing a song they had heard a parent sing before may fulfill a social function, with survival value. Specifically, new individuals who convey information previously transmitted by the parent are likely to be members of the in-group, to whom one can turn for survival needs if the parent becomes unavailable.

Remarkably, the infants in Mehr and colleagues' (2016) study also retained the short song well enough to discriminate between the familiar song and the other melody with the same lyrics and rhythm after a period of eight whole months had passed without listening to the song. The striking findings of studies reported in this section on infant memory for music reflect the importance of employing procedures that span longer time periods and involve home settings, to complement findings from infant studies conducted through single exposures in a laboratory setting.

Summary

So far, we have seen that infants are sensitive to elements of music such as pitch, consonance and dissonance, melody, tonality, rhythm, and meter, and that they appear to recognize familiar music after a time delay. Researchers are studying many other topics, including infants' perception of *timbre*, *phrase structure*, *harmony*, and *polyphony* (e.g., Céline & Trainor, 2013; Jusczyk & Krumhansl, 1993; Trainor, Lee, & Bosnyak, 2011), to name just a few. As with the research involving prenatal procedures, research findings in infant studies must be interpreted with a measure of caution. The innovative methods that have been devised for infant auditory research involve making some inferences about infants' perceptual abilities from indirect measures such as looking time, changes in heart rate, and neural responses. Nevertheless, when findings from multiple studies using different methods and stimuli converge, we can draw some (measured) conclusions about infants' apparent sensitivities to particular parameters of music that would otherwise remain mysterious and unanswered.

Infant-directed speech and singing

In the final section of this chapter, we expand the scope of discussion to the social contexts of infants' lives. Specifically, we explore caregiver–infant interactions involving expressive speech and singing directed to infants.

Infant-directed speech

Infant-directed (ID) speech refers to the style of speech that many speakers use to communicate with infants. The distinctive features include higher pitch, greater fluctuations in pitch and loudness, and fewer syllables than speech directed to adults (Broesch & Bryant, 2015). Thus, it differs from ***adult-directed (AD) speech*** with respect to ***prosody***, which refers to the 'musical' qualities of speech such as the melody, rhythm, tempo, and intonation of spoken language. ID speech can also be differentiated from other expressive styles of speaking, such as speaking to one's pets: for instance, English-speaking adults use a higher voice, express more emotion, and hyperarticulate their vowels to a greater degree in infant-directed than in pet-directed speech (Burnham, Kitamura, & Vollmer-Conna, 2002).

Pitch contour seems to play a critical role in conveying the speaker's intention, as 'the melody carries the message in speech addressed to infants to a much greater extent than in speech addressed to adults' (Fernald, 1989, p. 1505). Common patterns in the 'melody' of ID speech include rising contours for eliciting infant attention or encouraging the infant to vocalize or respond, falling contours to soothe an infant or end a turn-taking sequence, and bell-shaped contours for maintaining the infant's interest or conveying approval (e.g., Papoušek, Papoušek, & Symmes, 1991). These pitch contours are used quite similarly in different languages, including tone languages (in which pitch inflection conveys meaning, such as Mandarin Chinese) and non-tone languages (such as English and German), although some cultural variations have been observed (Saint-Georges et al., 2013).

Many studies show that infants prefer infant-directed speech to adult-directed speech, as indicated by measures of attention such as length of attentive looking time. In fact, infants show a preference for ID speech even if it is not in their native language. For instance, Werker, Pegg, and McLeod (1994) found that infants preferred ID Cantonese speech to AD Cantonese speech regardless of whether their own native language was Cantonese or English. Infants who wear hearing aids also develop a preference for ID speech, though it emerges somewhat later than the norm, and infants with cochlear implants (see chapter 3) show a preference after 9 to 12 months of hearing experience (Bergeson, 2012). One reason for the salience of infant-directed speech is its heightened emotionality. Neuroscientific research has shown that changes in EEG responses correlate with the emotionality of infant-directed speech among 9-month-old infants (Santesso, Schmidt, & Trainor, 2007).

The distinctive qualities of ID speech capture attention and facilitate greater comprehension, as shown by 21-month-old infants' enhanced ability to learn labels for new objects in sentences presented in ID speech but not in AD speech (Ma, Golinkoff, Houston, & Hirsh-Pasek, 2011). Further, ID speech promotes social synchrony between caregiver and infant, and supports joint attention and turn-taking (Gratier et al., 2015; Roberts et al., 2013). Signals such as rising contours to initiate dialogue and falling contours to terminate them facilitate smooth turn-taking between partners, and regulate the timing of events within an interaction. As Custodero (2002) has described it: 'The musical nature of our coordinated conversations with babies is similar to the musical conversations between jazz musicians. Both types of communication require a shared understanding of the musical structure, like "your turn, my turn"' (p. 6).

Infant-directed singing

Many people also adopt a distinctive style when singing to an infant. ***Infant-directed singing*** is characterized as having higher pitch, greater dynamic range, slower tempo, and longer pauses between phrases (Trehub, Hill, & Kamenetsky, 1997) than singing that is not directed to infants. Although most studies have focused on mothers, fathers also adopt this style of singing to their infants, though they do so less frequently and do not raise their vocal pitch in ID singing as much as mothers do (O'Neill, Trainor, & Trehub, 2001). Even siblings under the age of 3 years sing at a higher pitch and use an animated 'smiling' voice when singing to their infant sisters and brothers (Trehub, Unyk, & Henderson, 1994).

Just like ID versus AD speech, infants show a strong preference for ID singing over other styles of singing. Interestingly, even when caregivers are instructed to do their best to sing

Figure 8.5 Infants often seem mesmerized by the sound of singing directed to them. Six-month-old Hazel is transfixed as her mother sings 'Sing, Sing a Song' to her.

Source: Photograph used with permission from the parents. Copyright © Arianne Abela.

as if they were singing to their infants, infants can discriminate between infant-present and infant-absent singing styles, showing a clear preference for infant-present singing. For instance, Trainor (1996) found that 5- to 7-month-olds looked significantly longer at a loudspeaker that played infant-present recordings of a song than at the one that played infant-absent renditions. Mothers' and fathers' singing in the presence of a real infant is also judged by adults to be higher in pitch, slower in tempo, and more engaging in vocal quality ('warm voice,' 'smiling sound') than when doing their best to simply simulate singing to an infant (Trehub et al., 1997). Facial gestures and body movements such as smiling and swaying (that occur naturally in the presence of an infant) may affect the vocal quality of singing. Smiling, for instance, alters the shape of the vocal tract, and the acoustic quality of the singing voice takes on a warmer timbre (Sundberg, 1982).

Music seems to play a key role in coordinating the social interplay between a parent and infant. One-year-old infants maintain longer physical contact and eye contact with parents, engage in longer sequences of interaction, and express more positive emotions during play that involves music (such as singing, and musical toys) than during nonmusical play (Mualem & Klein, 2013). The structure of the song may organize their interactions during episodes of ID singing. For instance, Longhi (2009) found that mothers intuitively emphasize the structure of music in their ID singing by exaggerating features such as the metrical and phrase structure of a song (in much the same way that speakers may exaggerate prosody to convey the structure of an utterance in ID speech). For instance, mothers convey the temporal structure of a song 'like an orchestra conductor' (p. 207) by emphasizing upbeats (the weaker beats) and often extending these upbeats in duration in order to give a clear signal before a downbeat. In turn, infants synchronize their body movements with their mothers more frequently to the upbeats and downbeats that are emphasized by their adult partner's singing than to other moments within the song.

When it comes to vocalizing together, parents and infants can literally be 'in tune' with each other. Van Puyvelde and colleagues (2010) studied free-play interactions of mother–infant pairs and found that during times of mutual vocalizing, they often adjusted their voices to each other so that their pitches became consonant with each other, a phenomenon they refer to as 'tonal synchrony.' When the researchers analyzed the tones sung by the pairs, most intervals could be expressed by simple ratios such as 2:1 (octave), 3:2 (perfect fifth), and 5:4 (major third), which sound pleasing or consonant. In a subsequent study, van Puyvelde and colleagues (2013) conducted moment-by-moment micro-level analyses of 854 vocalization episodes of mothers and their 3-month-old infants. Interestingly, they found that periods of such 'tonal synchrony' co-occurred with moments during which the pair had just experienced mismatched social engagement, but were reengaging with each other. Conversely, dissonant vocal exchanges co-occurred with prolonged periods of mismatched social engagement.

Real-world applications

The effects of ID singing on attention, arousal, and coordinating actions of the infant make singing a useful tool in therapeutic contexts, for instance with infants of depressed mothers. Depressed mothers often lack the facial and vocal expressivity to sustain infants' attention for positive mutual engagement. Many depressed mothers also have difficulty recognizing and attending to the emotions of their infants. Even their ID singing often lacks social coordination, as depressed mothers tend to sing at a faster tempo than non-depressed mothers (de l'Etoile & Leider, 2011), leading to a singing style that is emotionally less expressive and less responsive to the infants' reactions. When combined with other interventions, ID singing can be used as a tool to help depressed mothers, through coaching mothers to be more emotionally 'in tune' by singing more expressively and imitating the responses of the infants during singing episodes (de l'Etoile, 2006). Exposure to only 10 minutes of ID singing has also been shown to modulate the arousal level in infants in the direction of balanced arousal (Shenfield, Trehub, & Nakata, 2003), which is helpful as infants of mothers with depression often have difficulty with self-regulation.

In daily interactions with infants in nonclinical settings, singing is also put to practical use. Because singing helps regulate infants' emotions or arousal levels (Shenfield et al., 2003), it is helpful for eliciting attention, inviting interaction and play, calming restless infants for feeding and changing, soothing, and inducing sleep. ID singing may also strengthen the caregiver–infant bond by maintaining infant attention for sustained interaction, promoting emotional synchrony between caregiver and child, and eliciting positive affect in the infant, which in turn can elicit further engagement by the caregiver (Dissanayake, 2000). Parental singing often structures daily routines such as mealtimes, bedtimes, and chores at home (Custodero, 2006) and becomes increasingly interactive as the child becomes a participant, as we shall see in the next chapter. In all these ways, infant-directed singing may have adaptive value for children and their caregivers.

Coda

This chapter focused on the perception and cognition of sound and music, before and after birth. Far from the Jamesian notion of the newborn's experience of the world as 'one great blooming, buzzing confusion,' we see an emerging portrait of a competent infant – whose auditory functioning begins a few months before birth. In their first year, infants already demonstrate sensitivity to the basic elements of music such as pitch, melody, tonality, rhythm, and meter. Moreover, infants do not only perceive and respond to music; they also create it. In the next chapter, we examine infants' and young children's engagement in musical behaviors – singing, moving to music, playing with musical toys and instruments – and examine examples of musical inventions and creations of their own making.

Note

1 We have adhered to the revised definition of the term 'pregnancy' by the American College of Obstetricians and Gynecologists (2013), which defines 'fullterm' pregnancy as 39 weeks 0 days to 40 weeks 6 days, and counts gestational age from the first day of the mother's last menstrual period.

9 Early musical development

It is often said that '*play is a child's work*,' as play provides a rich context for learning essential skills. Far from being a meaningless activity to while away the time, developmentalists view play as providing spontaneous learning opportunities that are significant to every domain of a child's life: physical, social, emotional, and intellectual. In this chapter, we will see how children's spontaneous responses to music and various forms of musical play lay the foundation for musical abilities that can be further supported through training and experience.

This chapter opens with an overview of developmental milestones in singing and movement to music in infants and young children. In this area, Helmut Moog (1976b) conducted one of the earliest investigations of the emergence of musical behaviors, involving over 8000 tests and observations of almost 500 infants and young children. Although subsequent observational and experimental studies have filled out and refined the details (as reflected in this chapter), Moog's work still provides the general frame for what we have learned about naturally unfolding musical behaviors during infancy and early childhood. The second section focuses on young children's exploratory play with musical instruments, and the development of children's improvisations and compositions.

Whereas the studies reviewed in preceding chapters focused on experimental studies in laboratory settings where variables can be carefully controlled, in this chapter we expand our focus also to include descriptive work and observational studies in naturalistic settings in order to capture children's exploration and engagement with music in real-world contexts – such as their homes, classrooms, and playgrounds. Laboratory and naturalistic approaches each have their strengths and limitations, and studies employing a variety of methods can lead to a richer understanding of the complexities of musical development.

Early singing and movement to music

Emergence of song

Our previous chapter described infants' responsiveness to singing. However, infants are not simply observers or recipients of music; they naturally participate in musical interactions with their own vocalizations. The earliest responses to music include cooing and babbling to the sound of singing or other music, which often emerges before 6 months of age (Moog, 1976b). In language development, 'babbling' refers to vocalizations comprising repeated consonant and vowel pairings, such as 'pababababa.' ***Musical babbling*** refers to similar infant vocalizations that are elicited during or immediately after the presentation of music (Moog, 1976b). In one study, Western infants' vocalizations in response to another person's singing voice were found to gravitate toward important components of Western scales, especially approximating (roughly) the first and third tones of the major or minor scale (Reigado, Rocha, & Rodrigues, 2011). Vocalizations by infants in response to hearing rhythmic speech or poetry, on the other hand, did not show these tendencies.

Songs children create

By around 2 or 3 years, many children invent ***spontaneous songs***, or improvisations of their own (Moog, 1976b). Figure 9.1 shows an example of 25-month-old Freya's vocal improvisations while playing with a doll, transcribed by author ST. Freya's song embodies many typical characteristics of improvised songs by 2-year-old children, as summarized by Dowling (1984). For example, her song is constructed out of short phrases repeated many times with little development or extension of ideas. Our notation of Freya's song is approximate, as many pitches were microtonal (i.e., they would fall in between the black and white keys of a piano), and the endings of phrases often switched to a speech-like tone. Most phrases were sung in one breath with the same general contour, rising and falling with the natural arc of a breath. There was no clear feeling of meter, as each phrase was sung in one breath with frequent pauses between breaths, though her singing reflected a fairly steady beat within the phrases (Dowling, 1984, p. 145). Freya's 'performance' was expressive and accompanied by actions, 'rolling' (rocking) the doll and standing up and singing loudly at the 'stand up!' lyric. She included a graceful song ending: The last phrase is the only one that features a rising contour and a sustained (rather than clipped) final tone.

Figure 9.1 Spontaneous song by 2-year-old Freya. The broken staves represent her tendency to sing in phrases rather than metrically, and parentheses denote speech-like tones.

Source: Recording obtained by Amanda McFarland and transcribed by Siu-Lan Tan, used with permission of the child's parents.

Sole (2017) found that a fruitful time to observe spontaneous singing is shortly after parents leave a child alone in the crib to sleep at night. She documented a wide variety of songs in children aged 18 months to 3 years, including babbling songs, chants, spontaneous songs, and learned songs. Sole conjectured that these crib songs fulfill diverse functions, including self-soothing, practicing newly acquired skills (such as repeating new words), and processing transitions (e.g., singing lyrics like 'It's bedtime now'). Some studies have shown that children sing more often and more proficiently in their homes than in other environments (e.g., Trehub & Gudmundsdottir, 2015), and thus a growing number of researchers are studying singing in home settings.

Melody is not always the most prominent feature of children's early improvised songs. Words and rhythm are more distinctive in children's ***chants***, which are narrower in range of pitches than singing. Indeed, a cross-cultural study conducted by Bjørkvold (1992) in preschools in Norway, Russia, and the United States found repetitive chanting to be the most common form of group song among children. Moorhead and Pond (1941) found that preschool children often chant during solitary or social play episodes, when engaged in repetitive motor activities such as stacking blocks or jumping or running, and when giving commentaries of events (e.g., announcing a sudden outbreak of rain). Some episodes of solitary chant were observed, but children most frequently improvised together, demonstrating a surprising degree of social organization:

Child A: 'Red, red, gingerbread.'
Child B: 'Red, red, go to bed / Red, red, go to bed.'
Child D: 'Stand on your head / With gingerbread'.
(p. 14)

Along with songs and chants of their own making, Moog observed the emergence of what he called 'potpourri songs' by the ages of 2 or 3 years (1976b, p. 100). As 'potpourri' refers to a mixture of different items, ***potpourri songs*** combine spontaneous singing with fragments of words, rhythms, or melodies of conventional songs. We observed a lively example of a potpourri song by 2-year-old Freya, as shown in Figure 9.2. During snack time, she was proudly demonstrating her newly acquired skill of eating with a fork. She repeatedly sang an original phrase, 'Mama come here watch Rah [Freya] do it,' to the tune of 'London Bridge Is Falling Down,' with variations of her own that resemble an introduction and coda. Whereas other researchers focused on melodic and rhythmic qualities, Bjørkvold (1992) focused on the *functions* of children's song inventions, and viewed them as a form of purposeful communication. In this light, Freya's potpourri song may be framed not simply as a light ditty, but as a 'plea' to draw her mother's attention to her latest achievement of handling a fork.

Songs children learn

The ability to sing songs from a standard repertoire improves gradually in the early years. Trehub & Gudmundsdottir (2015) found that mothers play an important role in supporting their children's singing by intuitively providing frameworks to support an increasing level of participation and engagement in songs. For instance, the researchers found that mothers often invite infants to participate by encouraging them to mimic pantomime actions for a song (such as 'I'm a Little Teapot') and by stopping at the end of song phrases so toddlers can 'fill in' the last pitch and lyric (such as 'Tip me over and pour me ____'). Mothers gradually lengthen the gaps for children to interject parts of the song until they

Figure 9.2 'Potpourri' song by 2-year-old Freya. No time signature is included as roughly equal emphasis was placed on each note; bar lines have been added only for convenience of reading.

Source: Recording obtained by Amanda McFarland and transcribed by Siu-Lan Tan, used with permission of the child's parents.

can reproduce the whole song. In this way, mothers often serve as their children's first 'singing mentors.'

Moog (1976b) found that when songs were sung to 1-year-olds, very few (only 6%) were able to give rough imitations of song fragments, and by about 2 years of age, most children could only reproduce a general outline or rough impression of the song. The children tended to imitate words first, then rhythm (supported by the lyrics), and finally melodic contour. By 3 years of age, Moog found that most of the children could imitate a short, simple song although not very precisely. Words were still the most accurately reproduced feature (as shown in subsequent studies, e.g., Welch, 1998), and rhythmic errors were often made if the rhythm did not fit the syllabic pattern of the lyrics. By the age of 4 or 5 years, most children know a small repertoire of conventional songs, though the pitches and rhythms are not yet very precise. The sense of tonality is not yet stable, so children often shift into different keys between phrases until 5 years or older (e.g., Shuter-Dyson & Gabriel, 1981).

By 5 years old, many of the basic skills of song-singing are in place: Although each interval may not always be precise, children's singing reflects a clearer sense of tonal center across phrase boundaries, an underlying pulse is apparent, and the ability to sing with expression is acquired (Davidson, McKernon, & Gardner, 1981; Shuter-Dyson & Gabriel, 1981). For instance, when asked to sing songs to make an experimenter happy or sad, girls and boys as young as 4 years old sang the happy versions faster, louder, and at a higher pitch than sad versions (Adachi & Trehub, 1998). They sang in this expressive manner when reproducing familiar nursery rhymes, and also when creating their own melodies and lyrics in happy and sad songs that they improvised. Children's invented songs also increasingly convey conventional features of the cultural repertoire, such as repetition of ideas, four-phrase form, and a sense of closure in Western 5- to 7-year-old children's vocal improvisations (Davies, 1994).

Singing accuracy

Singing in tune (often referred to as '***singing accuracy***' in research studies) is a difficult task. It requires the complex coordination of internal feedback mechanisms and auditory feedback, such as the ability to adjust the posture of the vocal cords to the sound of one's voice (see chapter 11). According to Rutkowski (1997), young children typically begin with a chant-like approach that often focuses on the lyrics. Next, they sing within a narrow pitch range resembling the speaking register, and then move through a phase in which they alternate between speaking and singing register, before finally consistently using a wider singing range (see also Welch, 1998). It is possible that the initial narrow pitch range and 'words first' strategy are not general developmental trends *per se*, but may vary with the characteristics of the song and the context in which the child may be learning it (Trehub & Gudmundsdottir, 2015). Wide intervals or leaps continue to pose a challenge to young children as well as adults. For instance, both children and musically untrained adults often compress the octave leap in the third phrase of the 'Happy Birthday' song (Davidson, 1994), singing a smaller interval, and consequently sliding into a new key!

With respect to methods for strengthening singing accuracy, Rutkowski (2015) recommends teachers train children's ability to expand the range of pitches they can use. This strategy, she argues, facilitates greater accuracy when imitating pitches. Providing children with a vocal model (especially a voice similar to their own) has also been shown to increase accurate pitch-matching in singing instruction of elementary school children (Green, 1990). Children's singing accuracy is also strengthened when they can respond to real-time visual feedback of their pitch accuracy. For instance, computer programs that visually represent the target pitch and the pitch a child is singing are more effective for improving pitch accuracy than verbal feedback alone (Hoppe, Sadakata, & Desain, 2006; Paney & Kay, 2015). In addition, a six-year longitudinal study in Italy found that expectant mothers who attended weekly music activities during the last three months before giving birth, and later attended with their infants, were more likely to have children who could sing in tune at 2 to 3 years of age (Tafuri, 2008; see also Gerry, Unrau, & Trainor, 2012). Thus, it appears that early engagement in music may influence the progression of singing accuracy.

Indeed, singing experience seems to be important to cultivate pitch accuracy throughout the lifespan. Demorest & Pfordresher (2015) compared data sets of three age groups (American kindergartners, middle school children, and musically untrained college students), and found that middle school children performed better than kindergartners on pitch accuracy in their singing. However, pitch accuracy dramatically *reversed* for college students, so that their performance was similar to the kindergartners. The researchers speculated that the decline in accuracy was not due to general developmental factors, but to the fact that general musical instruction ceases in most US middle schools. Thus, unless students at this age (i.e., 11 to 12 years) seek singing activities for themselves, singing practice usually stops, and therefore singing accuracy declines.

The ability to sing a common repertoire of learned songs leads to the possibility of singing in synchrony with groups, so that singing takes on an increasingly social character. As one develops the ability to match pitches and coordinate timing with others, there is often a shift from solitary singing to more collaborative creations, and enculturation into group ensemble singing. Indeed, an association has been found between well-developed singing skills and a positive sense of self and social integration among primary school children, which may stem from the typically social and collaborative nature of singing, which involves working toward shared goals, and matching and synchronizing behaviors with others (see Welch, Himonides, Saunders, Papageorgi, & Sarazin, 2014). Another study showed that engagement in group singing leads to more cooperative behavior in children than group art activities or competitive games (Good & Russo, 2016). These benefits are not limited to children, as group singing has been found to reduce stress and induce social flow or optimal interaction in adults (Keeler et al., 2015).

The learning of conventional songs marks the transition from a focus on improvisational and creative aspects of song to the acquisition of a particular musical repertoire that is taught or assimilated from the environment, including through media exposure. The child loses a measure of flexibility and originality in his or her song production, but in assimilating the values and conventions of a specific repertoire, becomes a more specialized member of a particular musical community.

Movement to music

Compared to the research on the emergence of singing, far fewer studies have focused on the development of movement to music. From 4 to 6 months of age, many infants respond to music by turning toward the sound and often ceasing movement, as if listening with rapt attention. After about 6 months of age, music begins to elicit increased movement, often accompanied by vocalizations (Moog, 1976a). Further, the movements infants make are more rhythmic when hearing music or regular drumbeats than when listening to audio recordings of adult-directed or infant-directed speech (Zentner & Eerola, 2010; Ilari, 2015).

Preference and responsiveness to certain meters over others seem to be shaped by experience. For instance, 7-month-old infants enrolled in *Kindermusik* classes showed a duple-meter bias (listening longer and more attentively to rhythms in duple than triple time) compared to a control group (Gerry, Faux, & Trainor, 2010). As the music employed in these classes is mostly in duple time, this preference seems to reflect their greater familiarity with engaging in listening – and especially moving – to music in this meter.

Spontaneous movements to music often emerge first in the head and neck, in the form of bobbing to rhythmic music. When infants can sit up securely and proprioceptive skills are advanced enough to make postural adjustments (around 7 months), they often bounce or sway their trunk to rhythmic music. Many infants also make rough ‘conducting’ motions to music for short bursts, though their movements are not yet synchronized with the music (Moog, 1976b). For instance, rousing music such as ‘McNamara’s Band’ elicited more bouncing movements in 9- to 13-month-old infants than Schumann’s slow and lyrical *Träumerei* (Trehub, 1993). However, these motions captured little more than the general pace of the music, and were not yet ‘in time’ with the music.

After achieving the motor milestones of standing (about 11 months) and walking (just before 12 months), infants begin to engage full-body movements in dance, stepping with alternating feet, bending the knees, and swinging their arms (as shown in Figure 9.3). Approaching 2 years of age, increasing strength and control of the feet and toes leads to the ability to tiptoe to music (Moog, 1976b). With increasing mobility and balance, children begin to use more space in their movements – hopping, running, and spinning to music.

The development of infants’ movement to music reflects two basic trends in human physical and motor development. There is the ***cephalocaudal trend*** (*head-to-foot*), or the tendency for muscular strength and control to proceed from the top of the body and progress downwards. This is seen in spontaneous movements to music that tend to begin with the engagement of muscles that control head-bobbing movements to music and progress downward to the engagement of the muscles in the trunk, legs, feet, and finally the toes – as described earlier. The other is the ***proximodistal trend*** (*near-to-far*), or the tendency for muscular strength and control to progress from the core of the body outward to the extremities. This is demonstrated in infants’ tendency to bounce and sway the trunk before beginning to engage the limbs, and finally the fingers and toes in their movements to music. Although change occurs rapidly in the infant and specific milestones may appear to be unrelated or

Figure 9.3 When Simon hears music, he often signs 'music' and then 'ear,' and begins to dance. At 21 months, his dancing includes full body movements, but it will be some years before he will be able to synchronize his movements to music.

Source: Photographs used with permission from the parents of the child. Copyright © Autumn Hostetter.

random when singly observed, it is interesting to note that, as a whole, motor responses to music progress in an orderly fashion.

Feeling the beat

The ability to move in synchrony with a beat takes a long time to master, and is rarely seen before 4 years of age (Drake, Jones, & Baruch, 2000). Simply beating or tapping time at regular intervals is difficult for most young children. Most 3-year-olds have difficulty keeping a regular beat when asked to tap the pulse of short pieces of music (Malbrán, 2000), and even when clapping in time to a metronome (Fitzpatrick, Schmidt, & Lockman, 1996). Further, the ability to engage the whole body in musical movement lags far behind the ability to move isolated parts of the body rhythmically. For instance, Rainbow (1981) observed over 150 preschool-aged children during a three-year longitudinal study on the development of rhythmic abilities. He found that 3- and 4-year-old children were more accurate when clapping a steady beat to music than when marching to the beat of music, and least accurate when both marching *and* clapping simultaneously to music.

One important limitation of children's ability to synchronize their movements to music is the range of tempos to which they can adapt. Children produce much faster spontaneous tempos than do adults. Specifically, the ***spontaneous tempo*** (or the rate at which tapping feels most comfortable) for 4-year-olds is approximately 150 beats per minute (e.g., Drake, Jones, & Baruch, 2000), compared to 100 beats per minute in adults (see chapter 6). In one study, 2- to 4-year-old children made spontaneous body movements roughly in time to music for short spurts if played at about 140 beats per minute, but did not adapt the tempo of their movements to match the music when played significantly faster or slower than this (Eerola, Luck, & Toiviainen, 2006).

Synchronization to a beat may be facilitated by co-action with another player. Kirschner & Tomasello (2009) asked children (who were 2.5 to 4.5 years old) to drum along with a human partner, a drumming machine, or a loudspeaker playing a drumming sound. At all ages, children were found to match the timing of the drumming more precisely when drumming together with another person (an adult researcher), and they were better able to adapt the tempo of their drumming away from their own natural tempo in order to synchronize with the beat when playing in the social condition. The results of this joint drumming study were particularly interesting because, as mentioned earlier, previous studies (e.g., Drake, Jones, & Baruch, 2000) showed that children under 4 years of age were unable to synchronize their actions to an external beat, suggesting a reduced capability for entrainment. Kirschner and Tomasello argue that this may be because previous studies had presented rhythmic tasks (such as tapping to a metronome or musical sequence) that lacked the social element inherent in much of music-making.

Over time, children's movements gradually reflect more parameters of music. Kohn and Eitan (2016) played musical excerpts that changed in pitch, loudness, or tempo, and asked children to move in a manner that would show a friend 'what the music is like.' Children aged 4 to 8 years old spontaneously expressed an increase in these parameters (i.e., rising pitch, *crescendo* or getting louder, *accelerando* or speeding up) with movements that increased in speed and energy, extended the limbs outward, and produced forward motions. In contrast, they conveyed a decrease in these parameters with less movement and smaller motions. A study by Boone and Cunningham (1998) showed that children are also able to decode some emotions conveyed by expressive dance movements, such as sadness, happiness, fear, and anger. For instance, when dancers exhibited a prolonged downward gaze and less muscle tension in the body, 4- to 8-year-old children labeled the gesture as 'sad.'

In sum: In the early development of movement to music, we see a progression from short bursts of activity to more sustained and controlled motor actions, and from infants' solo movements to young children's ability to coordinate their actions with partners. Further, children's spontaneous movements increasingly reflect the general structural features and expressive parameters of music, though not yet in synchrony with it. While we have discussed singing and movement separately, it should be noted that movement and vocalizations are often interconnected in infants' and young children's engagement with music.

Musical play and invention

As we shall see, innovative approaches to early childhood music education are attuned to the ways in which young children naturally integrate music into their daily lives, and further support and nurture these activities before more structured music instruction begins. The underpinnings of our discussion on music learning in young children are provided by the application of 'constructivist' theories of Piaget and Vygotsky to educational contexts. In both views, the child is driven by an inherent curiosity and actively 'constructs' his or her own knowledge, but the pathway to mastery takes somewhat different routes (as explained in the next sections). A full description of these complex theories lies beyond the scope of this chapter, so we will mainly focus on how these theories have been applied in educational contexts.

Exploration of sound

The active role of the learner is fundamental to Jean Piaget's (1896–1980) theory of cognitive development. His constructivist approach assumes that learning is driven by one's own curiosity and active exploration. The ***sensorimotor stage*** (the first of the four stages of cognitive development outlined by Piaget [1936], spanning from birth through 2 years) is a period of knowledge construction primarily through the use of one's senses and motor actions. For instance, through a series of increasingly complex 'circular reactions' (i.e., actions that are repeated many items, as they produce interesting consequences), sensorimotor exploration leads to the discovery of the consequences of motor actions and manipulation of objects. For instance, an infant who repeatedly hits, scrapes, and drops a spoon on a tray quickly discovers that different physical actions on objects can produce different timbres, pitches, and rhythms. Like 'little scientists' methodically repeating experiments, children are constantly building a knowledge base about cause and effect, physical laws, and physical properties of objects in their environment.

Just as the infant discovers the sound-making possibilities of a spoon by performing actions on it, young children in the ***preoperational stage*** (the second of Piaget's four-stage theory, corresponding with ages 2 to 6 years) construct knowledge through more complex forms of active exploration and builds some intuitive knowledge structures (i.e., 'schemes') as representations of their experience. Exploration becomes more methodical and organized, as young children systematically test different possibilities in their play with objects, and begin to anticipate consequences and adapt their actions.

Teachers who apply Piagetian principles to early childhood education view themselves as facilitators for children's active learning, by providing rich resources for hands-on exploration and setting up opportunities for making discoveries at one's own pace. The discussion that follows focuses on how to effectively support musical exploration during early childhood. In accordance with Piaget's ideas, two themes that will emerge in the following sections are the importance of self-directed musical play in developing the child's sense of *agency*, and the need to focus on *the process* (not just the product) of young children's early music-making.

Free musical play

A vibrant topic in both the practice and research in music education concerns the role of young children's 'free musical play' in music learning. In most traditional teaching approaches, the teacher directs and structures most learning activities. In ***free musical play***, however, the teacher provides the materials and rich opportunities for experimentation, and lets children structure their own learning through exploration, while being present to encourage their discoveries. Such a pedagogical approach may be seen as an application of Piaget's theory, as the idea of learning through *self-directed* active exploration and discovery is at its core.

In an early study, Gladys Moorhead and her colleagues observed 2- to 6-year-old children's free play with musical instruments at the Pillsbury Foundation School in California (Moorhead & Pond, 1942; Moorhead, Sandvik, & Wight, 1951). As a goal of the school's curriculum was to allow children to discover 'natural forms of musical expression,' emphasis was placed on young children's freedom to explore musical instruments with minimal adult intervention. No formal music class period was assigned, and children were free to enter a well-equipped music playroom to engage with a variety of instruments from around the world (gongs, cymbals, bamboo sticks, *maracas*, drums, and various string instruments) on their own at any time.

In line with Piaget's observations, Moorhead and colleagues discovered that young children's spontaneous explorations of musical instruments have a repetitive and methodical character, as if 'investigating' the possibilities of sound. When free to experiment on their own, children's activities were not chaotic, but quite systematic and orderly, and common

Figure 9.4 Free play allows Elle to explore and discover the sound-making possibilities of these musical instruments on her own.

Source: Photograph used with permission from the parents of the child. Copyright © Jessica Phillips-Silver.

patterns were observed in the way they explored sound in novel instruments. For instance, when examining a new instrument, most children appeared to be first drawn to *timbre*:

> *April 2.* Playing the Chinese theater drum with a hard-headed striker, Carl [3 years] discovered that the drum produced two different tones, one from the central stretched skin, one from the wide skin-covered wooden frame. He played a fast staccato beat, fairly steady, intent upon the effect produced by alternating the two tones. He then used the French bell series, alternating between the bells and the theater drum. Apparently disliking the harsh sound of the bells when struck by the hard striker he was using, he struck them seldom, continuing his interest in the two drum tones.
>
> (Moorhead et al., 1951, p. 95)

Also emerging early during initial exploration was experimentation with *variations in loudness*, testing the soft and loud sounds an instrument can make. Consistent with Piaget's view, children appeared to learn about dynamic range by connecting the various specific actions they performed (such as forceful versus light motions) with their expressive consequences. Experimentation with variations in *tempo* (speed) and *rhythm* emerged later, after the child was already accustomed to the instrument, as seen in Roy's play with a small drum:

> *May 12.* Roy [4 years] hung a small [Native American] Indian tom-tom around his neck and walked around with regular steps, drumming with two sticks in simultaneous beat,

> the beat coinciding with the right foot; almost without accent. He now began to drum in three-eighths time, using two sticks simultaneously .... He accelerated his tempo continuing for a time the effort to make his feet and drum beat coincide ....
>
> (Moorhead et al., 1951, p. 98)

Thus, during free musical play, children's curiosity often leads to spontaneous discoveries reached by their own initiative: Perhaps most importantly, free musical play allows children the time to become fully engrossed in a self-directed activity. Traditional teacher-led classroom activities are often changed with new directions every few minutes in order to hold the interest of the class, providing few opportunities for deep absorption in independent activity. Under these conditions, exploratory play is often prematurely terminated when the teacher directs the class to a new activity (Berger & Cooper, 2003), fracturing the child's concentration and disrupting the process of learning and discovery. By interspersing structured musical activities with time for free play, as some early childhood methods do, young children learn to keep in step with some teacher-led exercises while also having some time to engage in their own musical play without intrusion.

Engaging young children in music through free play emphasizes the importance of introducing basic musical concepts by experiencing them through movement, singing and playing instruments, and listening. Learning to read musical notation is not typically introduced in many child-oriented methods until children have developed a general familiarity with music. Laying down the foundations of musicality before introducing notation may avoid producing what Schleuter (1997) refers to as 'button pushers,' who simply view notation as a prescription for what fingers to put down, without any real musical sensitivity for the sounds they are creating. This ***sound before symbol*** approach (which advocates for building a strong foundation of rich aural engagement with music before introducing staff notation) draws on a long history of philosophies of music education (see Abeles, Hoffer, & Klotman, 1994). Studies have shown that many conventions of standard Western notation are counterintuitive to beginners (see Tan, 2002; Tan, Wakefield, & Jeffries, 2009) and that fundamental aspects of representation of sound such as the direction in which symbols are to be read may be perplexing to individuals who have not acquired skills in writing and reading text (Athanasopoulos, Tan, & Moran, 2016).

Learning in social context

Whereas Piaget viewed children as independent explorers who construct knowledge from discoveries made largely on their own, Lev Semionovich Vygotsky (1896–1934) was more concerned with the social and cultural influences on children's cognitive growth in the rich context of daily social interactions with parents, teachers, and guided interactions with peers. Thus, a Vygotskian classroom emphasizes *assisted discovery* through teacher–student and guided student–student interactions. Vygotsky believed that 'learning awakens a variety of internal developmental processes that are able to operate only when the child is interacting with people in [the] environment and in cooperation with ... peers' (1978, p. 90).

A central concept of the theory is the 'zone of proximal development' (ZPD), which Vygotsky defined as 'the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers' (1978, p. 86). Within this zone lies the greatest potential for learning, because it contains potential abilities that are ready to develop, but may not emerge without the proper social guidance.

Drawing on Vygotsky's ideas, psychologist Jerome Bruner used the term ***scaffolding*** to encompass any strategies that teachers and others guiding the child use to provide a framework that supports the learner's still-forming or 'budding' skills to blossom within their ZPD (Wood, Bruner, & Ross, 1976; see also Rogoff, 1990). Just as the scaffolding at a construction site supports a new building as it is being completed, effective scaffolding provides a supportive structure as learners build new skills and knowledge.

Supporting musical exploration

What sort of scaffolding best supports young children's musical play and improvisation? Berger and Cooper (2003) observed children and their parents for 10 weeks as they participated in a musical play program that included free play with a variety of musical instruments. With respect to adult behaviors that *enhance* musical play, Berger and Cooper found that adults who (1) valued the child's activities and explorations (even if not 'correct' by societal standards), (2) were willing to follow the child's lead, and (3) provided encouragement (while refraining from directing the child too much) helped to sustain children's free play with musical instruments.

Indeed, the mere presence of an interested adult may be sufficient to enhance play in young children. In another study, 2- and 4-year-olds played with a xylophone either in a solitary fashion or in the presence of an adult. Young (2003b) found that the presence of an attentive and responsive adult sustained children's play for longer periods of time, and led to more inventive and complex explorations of the xylophone – even if the adult simply watched. Children spent the least time playing on a xylophone when alone (1 minute

Figure 9.5 Ryder's early encounter with a banjo. Some studies have shown that young children's free play with a new musical instrument is sustained when an adult is present but not overly directing the child.

Source: Photograph used with permission from the parents of the child. Copyright © Andrea Kent.

40 seconds on average). They played almost twice as long when attended by an unfamiliar adult who was musically trained (3 minutes 12 seconds), and by far for the longest time when attended by a familiar adult with no musical training (6 minutes 49 seconds).

By merely being attentive and interested observers, adults seemed to scaffold young children to engage in more sustained and complex forms of play with a musical instrument than during solo explorations. It is interesting that the presence of musically untrained adults seemed to encourage the lengthiest play. Musically trained adults may unintentionally constrain play by subtly encouraging and rewarding 'correct' ways of playing, thereby narrowing the range of exploratory behavior and diminishing the child's intrinsic curiosity.

The importance of child-led strategies in scaffolding children's musical play was further examined in a study by Young (2003a), who identified three styles of interaction that *hinder* or *stifle* children's interest in a turn-taking or matching game in a music class: (1) an adult taking a commandeering or directive role by telling children how to play instruments ('play it like this'); (2) an adult playing only his or her own melodies and rhythms, failing to match or incorporate the children's musical ideas; and (3) an adult engaging in poor turn-taking, such as starting to play too soon or too late to sustain or continue a group rhythm (p. 111). Similarly, Berger and Cooper (2003) found that an adult suddenly approaching and hovering over a child who was deeply absorbed in an activity, or correcting and directing a child's actions, tended to *extinguish* the child's play.

Extending the scope of observations to the home environment, Koops (2012) found that home videos of parent–child musical interactions yielded similar findings about what supports musical play in the home (i.e., child-led activities: 'Play this now, Dad,' 'Jump with me, Mom'; p. 23) and what behaviors tended to terminate musical play (e.g., an adult becoming too involved, or giving too many suggestions). Thus, dominant adult styles of interaction seem to have a similar inhibiting effect on young children's musical play in both classrooms and home settings.

In particular, the findings of the studies reviewed here point to the importance of the adult imitating or elaborating on the *child*'s ideas. Young (2003a) observed two types of imitation in collaborative musical play between preschool children and adult partners: 'sequential imitation' (one person plays; the other person listens and then imitates) and 'imitation to synchronize' (both play together while making an effort to match each other, by sending and reading each other's nonverbal cues). Both forms of synchrony are integrated into group improvisations among musicians, particularly in the jazz tradition.

In sum, the somewhat mixed findings about the effects of an adult on young children's musical play may be explained by the adult's style of involvement. Adults often tend to be directive with children, but many studies show that a responsive and reciprocal style is most effective in supporting the musical play of very young children. During the early exploratory phase, an unintrusive but responsive adult partner may facilitate more sustained, varied, and complex musical play, enabling young children to play at higher levels within their zones of proximal development than when playing alone or unnoticed. Of course, as children become more skilled and continue in their music instruction, the balance shifts, and teachers and parents take a more active role in order to guide the technique and precision involved in mastering instrumental or vocal skills.

Observing process: Contextualizing children's musical activities

Tarnowski (1999) points out that like any other form of play, musical play 'is an activity in which being engaged in a *process*, rather than achieving a final *product*, is the goal' (p. 27, emphases added). To the child, the real value is in the *doing*, not in the final outcome (see Figure 9.6). However, few adults would regard actions such as scrubbing the keys of a toy xylophone with the 'wrong' end of the mallet as instructive to musical learning. Such activities are likely to be regarded as mere annoyances, as adults are often impatient to hear more mature products of music-making:

> Adults are accustomed to recognize musical production of children according to arbitrary standards of their own which merely draw lines near the peak of the enormous body of their musical experience and production. Each adult draws this line to suit himself, rejecting most of what is real music for the child at his own level.
>
> (Moorhead & Pond, 1942, p. 32)

Researchers, too, tend to focus on the product of children's music-making, often without much attention to process. Children's spontaneous vocalizations or improvisations with instruments are typically recorded, transcribed, and analyzed in research studies. However, young children's improvisations do not naturally fit the pitches, rhythms, or text of adult forms of notation; transcribing the sounds may therefore impose a structure or organization that is not found in the original performance. More importantly, much richness is lost when children's improvisations are analyzed without examining how the sounds are made, or considering the full context in which the vocal or instrumental play took place.

Figure 9.6 Emma's first attempt at music performance at 14 months of age.

Source: Photograph used with permission from the parents of the child.

The strong focus on product has also led to a lack of attention to the *kinesthetic* and *social* aspects of young children's music-making, such as the engagement of the body in the activity and in coordination with others. This was reflected in the aforementioned finding that young children synchronize better with a drumbeat when co-drumming with another person rather than a machine (Kirschner & Tomasello, 2009).

Playground handclapping songs serve as another illustrative case in point. Playground songs were once considered by ethnomusicologists to be rather simple in melody, rhythm, meter, and form. This assumption was based on analysis of transcriptions of melody and text, which are often in duple time (e.g., 2/4). However, when body movements and social coordination are taken into account, a more complex picture of this form of musical play emerges. For example, researchers have documented a popular three-beat handclapping pattern in playground songs of schoolchildren in many countries including United States, England, Australia, and Norway, which often bears a polymetric relationship to text that is usually in duple time (Marsh & Young, 2016). A standard three-beat pattern of clapping proceeds in this way:

Clap right hand face down, left hand face up, against partner's hands.
Both hands vertical to clap forward to meet partner's two hands.
Clap own hands together.

When combined with chants or songs in duple (two-beat) time, this three-beat clapping pattern forms a rather complex relationship to the lyrics, as it requires the children to feel or sense two rhythmic structures (i.e., duple and triple) at the same time. Handclapping songs based on intricate seven-beat and 13-beat patterns are also common among schoolchildren in the US and Australia (Marsh, 2008) and more complex polymetric patterns have been documented in many countries, such as duple-text/quintuple-clapping combinations in Portugal (Prim, 1989). In South Africa, handclapping songs involve the swift exchanging of roles that require the tight social and motor coordination of singing and action, incorporating elaborate call-and-response chants and songs, with improvised text and onomatopoeic elements, and dance elements engaging the whole body interspersed with complex clapping routines (Harrop-Allin, 2010). Much is missed if not observing the *process* of children's musical activities in fuller context – which is why naturalistic observation, case studies, and ethnographies are richly informative in this area.

As children enter the school years, musical play begins to take on a more social and collaborative character. Alongside their often structured formal classroom and/or private music lessons, an informal education is taking place in playgrounds and lunchrooms, and on school buses and within the home, as children coordinate chanting and singing with rhythmic body movements in handclapping songs, hopscotch, and jump-rope games, and refine ensemble timing with group chants and call-and-response games – all under the guise of 'play.' In these informal settings, transmission of songs and musical games takes place between peers, and especially from older children to younger children, so that multi-age school playgrounds offer a repertoire of options. This largely orally transmitted practice is fluid; variants of the songs and adaptations of singing games are constantly being co-created. As Marsh and Young (2016) state, 'In children's generative practices there is *no dichotomy* between process and product, so that the repertoire is constantly evolving' (p. 475, emphasis added).

Children's compositions

Children's musical compositions, too, can be most richly studied by examining both process and product. Kratus (1989) asked children aged 7, 9, and 11 years to compose on an electronic keyboard for 10 minutes. The 7-year-old children spent most of the allotted 10 minutes engaged in 'exploration' (playing new tone combinations and rhythms without repeating ideas). The 9-year-olds engaged in exploration, but also spent half the time engaged in 'repetition' and 'development' (variations of what they had played earlier in the same session). The oldest group, the 11-year-olds, divided the 10 minutes evenly between exploration, repetition, and development. Consequently, the youngest group's strategy for composing was similar to experimentation or improvisation, whereas the older children's compositions showed the beginning of basic form.

So when do children seem to shift their attention from the process of music-making to the products of their musical efforts? Based on the way they prioritized the allotted time for composition, Kratus (1989) concluded that the 7-year-olds were *process*-oriented in their approach to composition, as 'the process of exploring new sounds took precedence over the creation of a single composed product' (p. 19). The 11-year-olds, on the other hand, were more *product*-oriented, and 'seemed to be more focused on the task of creating a single product than on exploring new sounds' (p. 19).

Kratus' study employed a ***cross-sectional*** design (which compares the development of different children of different ages to infer the trajectory of development). Using a ***longitudinal*** design, Brophy (2005) followed the *same* 62 children for three years, from the age of 7 to 9 years. He also observed a shift toward increased musical organization (e.g., metric organization, repeated rhythmic motifs) in their instrumental improvisations, occurring most noticeably between ages 7 and 8 years. Brophy concluded that 'process orientation may be ending for most children around age 7, and product orientation may begin around age 8' (p. 131). Interestingly, we see a similar shift in other domains of development during middle childhood. For instance, in the area of art, Luquet (1913) and subsequent researchers observed that by around 7 or 8 years, children become preoccupied with conveying realism in their drawings, and begin to compare the images they create to an external standard.

Collaborative compositions

In contrast to early research focusing on the lone artist producing great works of art (Kozbelt, 2012), researchers are expanding the scope of study to social and collaborative aspects of composition by children. For example, MacDonald & Miell (2000) found that pairs of children (aged 10 or 11 years) who were close friends produced collaborative compositions that were rated by a 'blind' evaluator as being significantly better in quality than compositions by partners who were not previously familiar with each other. The compositions were rated on a set of scales designed to assess children's works of art, including ratings for how original, flowing, and vibrant they sounded.

Upon closer analysis, MacDonald and Miell found that the difference seemed to stem from the 'transactive' style of communication between friends, characterized by extending and elaborating on each other's musical ideas, giving suggestions, asking for clarifications and justifications for partners' contributions, and resolving conflicts, which facilitated the production of higher-quality compositions. This transactive style was apparent in the analysis of the partners' communication via both speech and music (e.g., the way they elaborated on each other's musical contributions by adding more tones), and was more beneficial to the process when instructions were open-ended rather than structured (see also MacDonald, Miell, & Mitchell, 2002).

Similar collaborative behaviors have been observed among much younger 4- and 5-year-old children engaged in group composition scaffolded by a teacher (St. John, 2006), including anticipation and expansion of teachers' and peers' ideas, and problem-solving. (For a discussion of scaffolding techniques for music composition based on Vygotskian techniques, see Wiggins and Espeland, 2012.) Notably, the studies we have cited here on collaborative creativity all draw on Vygotsky's sociocultural theory to highlight the role of joint action and social dialogue in shaping the creative process – elements that had often been overlooked in previous research on children's compositions.

Computer-based compositions

With access to increasingly elaborate but user-friendly digital tools, researchers are investigating how technology might facilitate the process of composition. Folkestad (1996) observed three main 'starting points' that adolescents (aged 15 and 16 years) used to embark on a composition: (1) begin with an impression of a musical idea; (2) select a genre or style of music; (3) or develop a musical idea that emerges while playing your musical instrument (a strategy that somewhat blends improvisation and composition). From this point, most adolescents proceeded with a horizontal or vertical orientation to the task. The 'horizontal' strategy involved first composing the entire melody, harmony, and overall form (with or without computer), and then using the computer to complete the arrangement and instrumentation to fill out the frame. The 'vertical' strategy involved using the computer from the start and working in sections, completing all layers before moving to the next section. Thus, in the vertical approach, composition and arrangement and orchestration were integrated rather than worked out in separate steps.

In a study focusing on much younger children (aged 6 to 8 years old), Nilsson and Folkestad (2005) tracked a group of children for 18 months as they engaged in computer-based composition. Each child had access to a synthesizer with a keyboard and a professional computer sequencer program, and was invited to create compositions inspired by images of nature scenes. The researchers found that providing digital tools allowed children without formal musical training to engage in composition, and that the compositions reflected form and structure (such as repetition and development of ideas) at an earlier age than observed by some aforementioned researchers such as Kratus (1989) and Brophy (2005), whose participants used only musical instruments. The music sequencer provided a portal to composing and arranging, allowing the children to layer instruments, repeat and mix motifs, and experiment with different tone and timbre combinations.

Indeed, other studies have shown that 'blind' evaluators could not distinguish between computer-based compositions created by 11-year-olds with two to four years of formal instrumental music tuition and those with no music training when assessing them on measures such as originality, form and organization, and technical goodness (Seddon & O'Neill, 2001). Similar findings were replicated in a subsequent study when evaluators judged computer-based compositions of adolescents (Seddon & O'Neill, 2006), suggesting that at the elementary stages of composition, musical creativity may simply be constrained by a lack of dexterity with instruments or lack of familiarity with musical concepts.

Interestingly, Nilsson and Folkestad (2005) also discovered through subsequent interviews that children who framed their activity as 'play' were more imaginative and engaged in their composition than those who interpreted it as a 'task.' This may be because the 'task' versus 'play' frame brings different ideas to the foreground, such as an orientation toward rules, conventions, and evaluation for tasks versus more of a spirit of exploration, originality, and imagination for play. Once again, we see how play facilitates learning, and echo the theme of this chapter: 'play is a child's work.'

Coda

Infants and children seem to show a natural inclination toward music through musical vocalizations, spontaneous singing, movements to music and rhythmic sounds, and the intrinsic curiosity they show for exploring sound-making objects and instruments. They are inventive in their engagement in music, naturally improvising and exploring creative facets of music-making. Thus, musical play programs and early childhood music methods may be viewed as simply *continuing* the trajectory of naturally occurring behaviors that children already engage in, and reinforcing the proclivity to music that they already possess. Child-oriented music methods sustain and cultivate basic skills that are already in place, providing a bridge to more structured instrumental or vocal lessons that will continue the trajectory of musical development toward more specialized skills and refined technique. The next two chapters on music practice and performance will focus on the development and refinement of these specialized musical skills as one continues along the pathway toward musical expertise.

10 Practice and musical expertise

'How lucky you are to be able to play that instrument.'
'It's funny – the more I practice, the luckier I get!'

The second line of the exchange above – 'the more I practice, the luckier I get' – is a popular quote that has been attributed to various professional golfers. However, in this context, it was quoted by Elizabeth Harré (former double bass player at the Sydney Opera House, and daughter of author RH), responding to an admirer. Whether observing a virtuoso string concerto or a masterful stroke on the putting green, observers often all too casually attribute a great performance to luck or inborn talent. But as anyone who excels in any field knows, the gap between innate capacities and mastery is filled with years of hard work and training!

While the last two chapters have focused on the emergence and development of basic musical abilities, we now shift our attention to the refinement of more specialized skills. Up to this point, we have shown how infants and children appear to have a natural propensity for music. However, specialization in the field requires much more than this. An essential element of this active involvement in music is the hours spent in individual or group music practice. The present chapter focuses on one of the most practical topics relevant to a musician's life: practice. Until quite recently, there was little research on music practice. Studies on music performance tended to focus on the product, rather than the process, of learning a musical work. It is only since the 1980s or so that a body of literature has emerged on this topic (Gruson, 1988). In this chapter, we review classic and current studies on music practice, and other topics relevant to the acquisition of musical expertise. We begin with a discussion of the relationship between deliberate practice and musical achievement, and the contribution of natural capacities or inborn 'talent.'

Practice and the acquisition of expertise

The role of talent

One of the central questions in the area of music practice is to what extent musical achievement is determined by inborn capacities or talent, and to what extent training and effort play a role? The definition of ***talent*** in general usage is not precise, although researchers tend to converge on five features proposed by Howe and colleagues, who clarified its meaning (though were not themselves proponents of this account of high levels of achievement):

> (1) [Talent] originates in genetically transmitted structures and hence is at least partly innate. (2) Its full effects may not be evident at an early stage, but there will be some

> advance indications, allowing trained people to identify the presence of talent before exceptional levels of mature performance have been demonstrated. (3) These early indications of talent provide a basis for predicting who is likely to excel. (4) Only a minority are talented, for if all [individuals] were, then there would be no way to predict or explain differential success. Finally (5), talents are relatively domain-specific [i.e., restricted to a specific area of skill, like music].
>
> (Howe, Davidson, & Sloboda, 1998, pp. 399–400)

It should be noted that practice is still assumed to play a role in talent-based views of achievement. Nobody claims that a talented musician should be able to sit down at the piano and immediately play a Rachmaninoff concerto! What differentiates talent-based accounts of musical achievement from views that focus more on the influence of training and experience is *how much* of an effect practice can have (cf. Ruthsatz, Detterman, Griscom & Cirullo, 2008), and *the degree to which* one is assumed to inherit the ability to benefit from practice.

The role of deliberate practice

In contrast to views emphasizing inborn talent, K. Anders Ericsson and colleagues argued in a seminal paper published in 1993 that the role of inborn capacities in achieving superior performance has been exaggerated or overemphasized in many different fields, including music (Ericsson, Krampe, & Tesch-Römer, 1993). At the time, the notion that superior performance is determined by innate capacities was rarely questioned, an idea that can be traced back to Sir Francis Galton (1869/1979), who observed that eminence in some fields tends to run in families. While not dismissing the role of genetic factors influencing practice and performance, Ericsson and colleagues argued that the contribution of genes to many skill-based domains has been greatly overemphasized, and that achievement is mainly a result of intense engagement in what they called 'deliberate practice.'

Ericsson and colleagues used the term ***deliberate practice*** to refer to highly structured practice with the explicit goal of improving performance, which involves tasks designed to address weaknesses and correct errors (Ericsson et al., 1993). In the domain of music, deliberate practice is exemplified by the hard work involved in improving skills by using techniques to correct mistakes and strengthen performance. Such techniques might include identifying concrete goals for each piece before beginning to practice, targeting errors and practicing the corrections intensively, slowing down fast passages to facilitate accuracy and gradually speeding up the tempo, vocalizing complex rhythms to improve precision, and memorizing passages in the score where one tends to stumble. It is not the same as ***mere repetition***, or simply repeating the same actions with the assumption that it will automatically lead to improvement.

Whereas mere repetition may be sufficient to bring about improvement in the early stages of learning a skill, progress is likely to be arrested before reaching higher levels of proficiency if relying mainly on drilling. Mental engagement in what one is doing appears to be key, which may be one reason why Carter and Grahn (2016) found that alternating between pieces frequently within a practice session (an 'interleaved' practice style) may be more effective than focusing on one piece at a time (a 'blocked' practice style). Shifting focus between pieces makes it less likely for the individual to fall into a state of thoughtless repetition. In some respects, mere repetition without mental concentration can even hinder progress. McPherson and Renwick (2001) found that children spend over 90% of their practice time simply repeating their pieces from beginning to end, without monitoring and correcting their mistakes.

Hallam (1995) found that 60% of 6- to 18-year-old string students (including advanced players) left uncorrected errors in their solo practice. However, unless conscious effort is made to improve performance and correct mistakes, errors may become 'practiced in' through repetition (Gruson, 1988).

Ericsson and colleagues devoted many years to examining how deliberate practice leads to the acquisition of expertise in a wide range of fields, such as sports, music, arts, and science. In the aforementioned widely cited review published in 1993, they found little evidence for genetic factors that served as reliable predictors for superior performance in these fields. Instead, they concluded from their review and own research studies that 'the level of performance an individual attains is *directly related to the amount of deliberate practice*' (Ericsson et al., 1993, p. 370, emphasis added).

Addressing the specific domain of music, Ericsson and colleagues (1993) studied three groups of violin students enrolled at a highly selective music conservatory in West Berlin: 'the best violinists,' a group of 'good violinists,' and a third group who met a lower standard of achievement. Estimates of hours of solo practice were computed by reconstructing year-by-year retrospective biographical data from each participant to estimate accumulated hours of practice over the years. Consistent with their claim, the researchers found that the 'best' violinists reported an average of about 7400 hours of solo deliberate practice by the age of 18 years, compared to 5300 hours for the 'good' violinists, and 3400 hours for the lowest-achieving group. Ericsson and colleagues also examined pianists at the same prestigious music conservatory, and found that 'expert' pianists reported an accumulated average of approximately 7600 hours of practice by age 18 years, compared to only about 1600 by amateurs.

On the basis of their research and review of the literature in many domains, Ericsson and colleagues concluded that the eventual level of achievement one reaches in a domain is not primarily determined by genetic factors with which some 'gifted' or 'talented' individuals are endowed, but by the greater amount of time engaged in deliberate practice. Put in bolder terms by Howe and colleagues (1998), talent may exist primarily as a cultural myth that fuels misconceptions about the value of training and experience.

Regardless of intensity or effectiveness, however, short-term engagement in deliberate practice is not sufficient to acquire expertise. Simon and Chase (1973) first discovered that very few players had reached the level of international chess master without at least 10 years of intense involvement in chess. The so-called ***10-year rule*** has since been supported by numerous studies examining eminent achievement in a variety of disciplines including mathematics, science, poetry, literature, tennis, swimming, long-distance running, and interpretation of X-rays (for a review, see Ericsson et al., 1993, p. 366). Some studies suggest that even more than 10 years of intense deliberate practice are required to achieve eminence in certain fields – including music. For instance, Sosniak (1985) found that concert pianists had studied piano for an average of 17 years before qualifying for the finals of their first major international piano competition, and Hayes (1981) and Simonton (1991) found that music composers had undertaken about 20 years of music study before producing their first eminent composition.

While effortful practice is the focus, the contribution of genetics is *not* overlooked in the deliberate practice view. For instance, Ericsson has proposed that 'the distinctive characteristics of elite performers are adaptations to extended and intense practice activities that selectively activate dormant genes that all healthy children's DNA contain' (2007, p. 4). In other words, intensive deliberate practice may trigger a cascade of biochemical changes that in turn stimulate cell growth and the transformation of cells. These low-level changes may eventually facilitate physiological adaptations necessary for superior achievement.

According to Ericsson's view, the genetic blueprint does not determine the upper limit of one's success, but rather establishes the building blocks for success that can be realized through training. However, as other researchers have pointed out, the effects of deliberate training alone may not be sufficient to account for exceptional abilities identified in very young children with only a few years of experience in their domain (e.g., see Ruthsatz, Ruthsatz, & Stephens, 2014; Winner, 2000). Further, some individual variables such as a child's temperament and level of persistence, which may contribute to why some practice longer and more effectively than others in the first place, may also stem from the genetic make-up of the child. For instance, a twin study has shown that a common set of genes may influence both musical discrimination abilities (i.e., acuity in detecting differences in pitch, rhythm, and melody) and a propensity toward practicing (Mosing, Pedersen, Madison, & Ullén, 2014).

Nevertheless, Ericsson and other proponents of the deliberate practice view maintain that the role of genetics and innate 'talent' has been overemphasized, and that training and effort (in the form of deliberate practice) play a far greater role in attaining expert-level musical achievement than previously assumed.

Beyond deliberate practice

Just as Ericsson pointed to the exaggerated importance placed on genetic factors in earlier theories of eminence, critics have argued that Ericsson and colleagues' emphasis on deliberate practice underestimates the role of other factors that contribute to achievement or mediate the effects of deliberate practice (Hambrick et al., 2014), such as genetic influences, the starting age of training, intelligence, and personality. As such, discussion of the contribution of talent to musical achievement has re-emerged.

Subsequent studies examining the link between amount of deliberate practice and musical achievement have produced mixed findings, some corroborating Ericsson and colleagues' findings (e.g., Sloboda, Davidson, Howe, & Moore, 1996; O'Neill, 1997b), and others failing to find a clear relationship (e.g., Duke, Simmons, & Cash, 2009; Madsen, 2004; Williamon & Valentine, 2000). Two ***meta-analyses*** (which use statistical procedures to quantify effect size based on many studies in a specific area) have also yielded different findings. One meta-analysis revealed that deliberate practice accounts for only about 30% of individual differences in music performance, leaving 70% of the variance to factors other than deliberate practice, including abilities influenced by genes (Hambrick et al., 2014). However, another meta-analysis that used more stringent criteria for inclusion of studies yielded findings that 'corroborate the central role of long-term deliberate practice for explaining expert performance in music' (Platz, Kopiez, Lehmann, & Wolf, 2014; see Figure 1 in that article for criteria). Ericsson has addressed the discrepant findings among studies by explaining that the original research on acquisition of expertise focused on highly accomplished experts, and that a strong relationship between deliberate practice and achievement may not be found in general samples of participants at sub-expert levels (see Ericsson, 2014).

In a sweeping review matching Ericsson and colleagues' broad scope of study in many fields of expertise, Brooke Macnamara and others examined the role of deliberate practice in 88 studies in the diverse fields of music, games, sports, and educational and professional achievement, and concluded that 'deliberate practice is important, but not as important as has been argued' (Macnamara, Hambrick, & Oswald, 2014, p. 1608). This 'important, but not as important' stance on deliberate practice leaves room to consider the contribution of other variables leading to musical achievement, including basic abilities that have been shown to be influenced by genes, such as working memory capacity (Hambrick & Meinz, 2011).

Although there is not yet much behavioral genetics research focusing on music ability, some studies point to a genetic basis for basic sensitivities to pitch and rhythm and sound patterns, and for memory for music, musical creativity (improvisation, arranging, composition), and singing ability, as well as rare abilities such as absolute pitch (see chapter 5) and deficits such as congenital amusia (or 'tone deafness,' see chapter 11) (for a review, see Tan, McPherson, Peretz, Berkovic, & Wilson, 2014[1]).

Some critics of the deliberate practice view argue that differences in innate capacities may explain why some people are able to reach higher levels of achievement in a domain than others after spending a comparable amount of time in intensive training. For example, an individual with greater working memory capacity or faster sensorimotor speed may gain more from devoting the *same* number of hours of deliberate practice to instrumental or voice training. Individual differences between people with respect to innate capacities may explain the great variability in the amount or duration of effortful training that has been found among elite performers in various fields. For instance, Gobet and Campitelli (2007) studied 104 chess players ranging from amateur to the most advanced grandmaster level, and found that some individuals who had devoted over 10,000 hours to playing chess had only reached an intermediate level. Further, among the chess players who reached the highest level, one elite player had only engaged in intensive chess playing for two years, whereas another had trained for 26 years.

An extensive analysis drawing from a twin database of over 800 twins also points to the close interplay of genes and environment in producing individual outcomes for music achievement. Specifically, the researchers found that approximately 38% of the variance in music practice itself is genetically determined (Hambrick & Tucker-Drob, 2015). In other words, one's genetic make-up may influence how much one is willing to practice and how one spends the practice time. Further, the study also showed that 'genetic effects on music accomplishment were most pronounced among those engaging in music practice, suggesting that genetic potentials for skilled performance are most fully expressed and fostered by practice' (p. 112). Again, it is important to note that for both Ericsson and critics who propose alternate views, the question about deliberate practice and genetics is *not if* – but exactly *how* – each contributes to musical achievement. This question endures in our field, its intricacies still hotly debated.

Expanding the scope of study

Building on the work of Ericsson and colleagues, researchers are seeking to understand the role of deliberate practice as one component embedded within a more expansive framework of individual and environmental factors shaping musical development and achievement. For instance, Bonneville-Roussy and Bouffard (2015) proposed an integrative model that accounts for how deliberate practice and motivational and self-regulatory mechanisms jointly influence musical achievement. This model is congruent with research showing that effective practice involves ***intrinsic motivation*** (as opposed to a focus on extrinsic goals such as physical rewards or approval of others), and that it is structured by ***self-regulation*** or the ability to monitor one's actions toward attaining goals (McPherson & Zimmerman, 2011; Schmidt, 2005). For example, one approach toward music practice that has yielded long-term positive effects was adapted from a sports psychology training program, and emphasizes such techniques as goal-setting, sharpening concentration, imagery, and confidence training through positive self-talk (Hatfield, 2016). Further, this integrative model is consistent with research pointing to the key role that teachers and parents play in supporting the learner's intrinsic motivation and self-regulation strategies during practice (e.g., Ericsson, 1996, p. 32; McPherson, 2009). Social support is an important factor in the acquisition of skills in any domain, as high achievement is rarely attained in isolation.

Other research suggests that personality factors may also influence the levels of practice reached in expert performers. Corrigall, Schellenberg, and Misura (2013) investigated what factors best predict the total years of lessons undertaken by a sample of 113 undergraduates. Their primary measure was the 'Big 5' personality inventory, which uses a large set of self-report items to classify personality types along a set of five dimensions: (1) openness to new experiences, (2) conscientiousness, (3) degree of extraversion, (4) agreeableness with others, and (5) neuroticism. Among these personality traits, ***openness to experience*** (i.e., how open one is to intellectual curiosity, new experiences, unconventional ideas, fantasy, aesthetic sensitivity, and exploring emotions) was the best predictor, even when important variables like socioeconomic status, IQ, and parental background were controlled using statistical procedures. Interestingly, the degree of openness to experience in one's parents also predicts years of practice (Corrigall & Schellenberg, 2015). Conscientiousness was higher than normal among well-practiced musicians, but did not predict years of practice independently from other variables, like IQ. Thus, openness stands out as being a distinctive, though not the only, predictor of music practice.

Taking a particularly inclusive perspective, Simonton (2014) has proposed an integrated approach to research on creative performance and expertise that encompasses both deliberate practice *and* talent-based frameworks. Simonton's approach involves the identification of relevant cognitive abilities (such as general intelligence, working memory, and spatial ability) and dispositional traits (such as motivation, personality traits, and values) as well as the genetic and environmental antecedents of these abilities and traits. Increasingly, proposed models and research programs are seeking to identify how deliberate practice operates – not in isolation, but in concert with many other factors to produce individual outcomes in musical achievement.

The role of informal practice

'Deliberate practice' usually applies to effortful time spent on ***formal practice*** (i.e., practice devoted to scales and pieces of repertoire assigned by one's teacher). In contrast, ***informal practice*** refers to time not spent on assigned exercises and pieces, but on 'playing for fun' (Sloboda et al., 1996). Informal practice includes playing old favorites that are no longer assigned, sight-reading music for enjoyment, playing by ear, engaging in improvising and other creative musical activities, and 'playing around' to experiment with sounds and motor patterns. Although informal practice may involve deliberate practice strategies such as self-correction and improvement over time, the goal is usually more immediate: to enjoy the activity.

Students are often discouraged from engaging in informal practice, and told to 'stop messing around and get back to your lesson.' However, informal practice targets and builds essential musical skills! For example, informal practice develops expressivity, as one is more likely to play 'with feeling' when less focused on the technical aspects of performance (Sloboda, 2005, p. 270). Activities such as playing by ear and improvising may also strengthen auditory imagery (i.e., the ability to imagine the sound of music without actually hearing it, as discussed later in this chapter), which in turn has been linked to other musical skills such as proficiency in sight-reading (Kopiez & Lee, 2006). Informal practice in groups, such as in jam sessions, hones listening and improvisation skills as one must respond quickly to musical ideas introduced by other players, and synchronize and harmonize with others. Further, playful involvement in music brought about by informal practice may sustain interest in music study, as there is evidence that students who engage in informal practice continue lessons longer than those who do not (McPherson & McCormick, 1999).

In addition, neuroscientific studies show that over time, practice strategies that develop aural (i.e., listening-based) skills may help performers anticipate what kind of pitches result from their actions, referred to as ***perception/action associations***. For example, Seppänen, Brattico, and Tervaniemi (2007) studied auditory processing in two groups of musicians. One group, referred to as the 'aural group,' reported that they improvised at least once a day, played by ear often or quite often, and practiced by listening to recordings of music they were learning. In other words, the 'aural group' engaged in informal practice as part of their practice routine. By contrast, the second group reported that they almost always performed from notation and did not improvise. Measurements of event-related potentials showed that the brains of musicians with aural training responded more strongly to changes in melodic structure (e.g., hearing one melody followed by a similar melody that contained an altered melodic interval) than the non-aural group. Thus, informal practice may also enhance the ability to perceive musical pitch relationships, even outside the context of a performance.

Surprisingly, people often seem to assume that playing by ear, improvising, and composing are 'gifts,' or skills that are naturally 'picked up' by the truly musical – and overlook that they also require practice! Individuals who claim to have no improvisation or composition skills or say they feel 'lost' without their sheet music have typically focused on formal practice, and

Figure 10.1 Jazz pianist Glenn Tucker spends two-thirds of his practice time on technical exercises and assigned repertoire (formal practice), and the rest on informal practice, mainly improvising at the piano. Tucker is a pianist at the Thelonious Monk Institute of Jazz Performance under the direction of Herbie Hancock.

Source: Photograph used with permission of Glenn Tucker and PKO Records. Copyright © John Lilley.

have devoted little time to playing by ear, improvising, arranging, or composing. Jazz pianist Glenn Tucker (shown in Figure 10.1) spends about two-thirds of his daily practice time on technical exercises, transcriptions, and classical repertoire, and the remaining third on informal practice, mainly playing standards by ear and improvising at the piano. Tucker states: ‘I think that one of the biggest goals in jazz is to be spontaneous, and the only way to be spontaneous is to practice being spontaneous’ (personal communication; see chapter 11 for further discussion). Johansen (2017) uses the term ‘explorational practice’ for such solo or group practice that integrates creative and improvisational aspects of musicianship into a regular practice routine, and emphasizes its role for a musician to find a personal ‘sound’ or ‘voice.’

In one of the few studies examining both formal and informal practice, Sloboda and colleagues (1996) examined the practice habits of music students with differing levels of achievement by studying practice logs and conducting interviews with over 250 students and their parents. By 17 years of age, the estimated total practice time for the best pianists was just under 7000 hours, which is in step with Ericsson and colleagues’ (1993) estimates for the ‘best’ violinists and ‘best’ pianists. In addition, more pertinent to this discussion, Sloboda and colleagues (1996) found that the top two groups also devoted more practice to *informal* practice than did the two lowest-achieving groups. In sum, formal and informal practice seem to develop different aspects of musicality, and involvement in both types of practice seems to produce the most well-rounded musicians.

Formalizing the informal

Woody and Lehmann (2010) found that musicians trained in styles such as jazz and pop outperformed classically trained musicians on ‘listen-then-imitate’ tasks, and attributed this to more time spent engaging in musical activities related to informal practice. Various innovative teaching methods such as the Ear-Playing Project (Varvarigou & Green, 2014) and ‘vernacular music’ approaches (O’Flynn, 2006) draw on methods that popular musicians (jazz, folk, pop, etc.) use to learn music, and integrate them as practice techniques in music lessons. Specifically, these teaching approaches emphasize learning new repertoire by listening to musical recordings and imitating what they hear, as popular musicians tend to do (as opposed to relying on notation). They incorporate student-selected music into the repertoire, emphasize practicing in groups as well as alone, and train improvisation and playing by ear from the very first lessons.

Informal elements can also be integrated into standard private lessons that teach notation and traditional repertoire. For instance, one of your authors (ST) taught private lessons for nine years as a certified piano teacher and regularly incorporated improvisation and ear-playing into lessons. She routinely spent five minutes at the end of lessons with young beginners and their parents, improvising together through singing and clapping games. She asked intermediate students to compose commercial jingles about their favorite products, and to improvise on melodic or rhythmic motifs. More advanced students were assigned to play and harmonize their favorite pop songs by ear. As more teachers continue to integrate ear-playing and improvisation into lessons alongside traditional scales and repertoire, the lines between formal and informal practice may begin to blur.

Practice strategies: A review of the research

Many students are not very deliberate or strategic in their practice. As one honest school-aged participant told researcher Gary McPherson: ‘If I’m doing really well I play it a few times.

If it's bad then I only play it once, except I know it should be the other way around' (2005, p. 18). In this section, we will discuss a few themes in the classic and current literature on music practice strategies: distributed versus massed practice, part versus whole approaches, mental versus physical practice, and aural models versus verbal instruction. Reflecting the existing literature, the focus will be on formal practice during solo sessions, although some principles may also apply to informal practice and group rehearsals.

Distributed versus massed practice

The issue of ***distributed versus massed practice*** focuses on the question: Is it better to divide practice time into many shorter sessions (distributed), or one single long session (massed)? Several studies show distributed practice to be more effective than massed practice for learning many types of motor skills (Oxendine, 1984). In an early study focusing on music practice, Grace Rubin-Rabson (1940a) found that massed practice leads to quicker – but not necessarily more precise – learning of short musical sequences. Students were able to learn music faster through intense sessions of massed practice. However, distributed practice led to better retention of the music after two weeks, and more accuracy when asked to transcribe the music from memory. Breaks between practice sessions give students an opportunity to take a fresh look at the music and to understand its structure more deeply. Distributed practice and frequent rest breaks during practice sessions are also strongly recommended for reducing the risk of developing various overuse injuries and other playing-related conditions, discussed later in this chapter.

Subsequent studies have shown that many advanced music students and professional musicians use distributed practice. For example, in Ericsson and colleagues' (1993) study described earlier, the 'best' violinists distributed their practice into three to four sessions per day that varied from about 30 minutes to two hours per session. They usually took two short breaks within each practice session, often playing for no more than about 50 minutes before taking the first break (reported in Krampe & Ericsson, 1995).

It is interesting to note that the top two groups of violinists in Ericsson and colleagues' conservatory study also napped more often than the less accomplished third group (1993, p. 375). This finding is in line with research on the critical role of sleep – independent of whether it occurs at daytime or nighttime – in *consolidating* memory for a sequential motor task (Fischer, Hallschmid, Elsner, & Born, 2002). ***Consolidation*** refers to the conversion of a memory trace that has just been acquired 'into a more stable form that becomes resistant to degradation' (Stickgold & Walker, 2005, p. 408). It appears that the memory trace for motor skill learning tasks continues to be reprocessed even during periods without intervening training, and that sleep plays a critical role in this 'slow component of learning' in adults (Maquet et al., 2003) and in children (Maski, 2015).

More recent research suggests that sleep helps to consolidate the effects of practice, but in a way that focuses on *conceptual* rather than *motoric* aspects of learning (Van Hedger, Hogstrom, Palmer, & Nusbaum, 2015). When pianists were tested after a 12-hour period with no intervening sleep, some forgetting was found. Pianists made more errors during recall than they did immediately after practicing. However, if pianists slept during a this span of time, no increase was found in pianists' number of *conceptual* errors, defined as playing a note that is in accordance with the key signature, but incorrect in terms of what is notated in the score. This advantage for sleep was not found, however, when measuring the number of *motor* errors pianists made, which occur when a pianist accidentally hits a nearby key that is unrelated to the music's key signature.

Part versus whole approach

Part versus whole pertains to the question of whether musical works should be divided into smaller segments during practicing, or practiced through as an intact whole. For example, Jennifer Mishra (2002) studied how intermediate students memorized a 36-measure étude in their own way, without being given any specific instructions on how to go about it. The following strategies spontaneously emerged:

Serial: Students using this approach tended to return to the beginning of the piece every time a mistake was made. There were usually many returns to the starting point, and the last part often underwent little practice.

Segmented: This strategy usually began by dividing a composition into segments, and learning each segment separately. The separate segments were then linked together and the joints between segments were carefully practiced.

Additive: Like the segmented approach, students employing this approach divided the composition into small pieces, but continually lengthened the pieces so that the composition was learned in rapidly expanding segments.

Holistic: This method entailed playing pieces from beginning to end, trying to avoid stopping and restarting. If a mistake was made, the student would backtrack just enough to correct the error, but not to the very beginning.

Out of the 80 instrumentalists in Mishra's (2002) study, the four who took the longest time (66–100 minutes) to memorize the 36-measure étude used Segmented or Serial approaches. The instrumentalists who took the shortest time (8–16 minutes) used Holistic or Additive approaches. Thus, the most efficient memorizers used methods that focused more on 'whole' and quickly expanding segments toward the entire piece, rather than 'part' learning (see also Gruson, 1988). However, it is possible that Holistic and Additive approaches may be superior only when learning and memorizing relatively short musical works. Case studies of musicians preparing much longer and more complex works have shown that they often rely on a combination of 'part' and 'whole' approaches (e.g., Chaffin, Imreh, & Crawford, 2002; Chaffin, Lisboa, Logan, & Begosh, 2010; Hallam, 1995).

In one study, Chaffin (2007) examined how concert pianist Gabriela Imreh memorized Debussy's *Clair de Lune* for public performance, by recording all her practice sessions from the first time she sat down at the piano with the piece until the day she performed it at a concert two weeks later. Figure 10.2 is a visual representation that conveys Imreh's practice activity during one practice session. The figure should be read *from bottom up*. Each line represents a continuous duration of playing; every time the pianist stopped and picked up again, this is depicted as a new line slightly above the previous one.

Figure 10.2 represents the pianist's first practice session. Imreh was familiar with *Clair de Lune*, but had never memorized it before or performed it publicly. The nearly continuous line at the bottom of the graph shows that she first 'ran through' the whole piece, with some brief hesitations shown by the broken lines. After getting a sense of the overall structure, she practiced Theme A (bars or measures 1–26), sometimes going back to the beginning or repeating shorts bits. She then practiced Theme B (bars 27–50) in a similar manner. From measure 51 to the end, progress was fast as the piece is in ABBA form. Imreh concluded the session by playing through the whole piece again, this time much more smoothly (as shown by very few breaks in the topmost line). With each session, the pianist relied less on the notation, and by the fourth session, Imreh was able to run through the whole piece without the score.

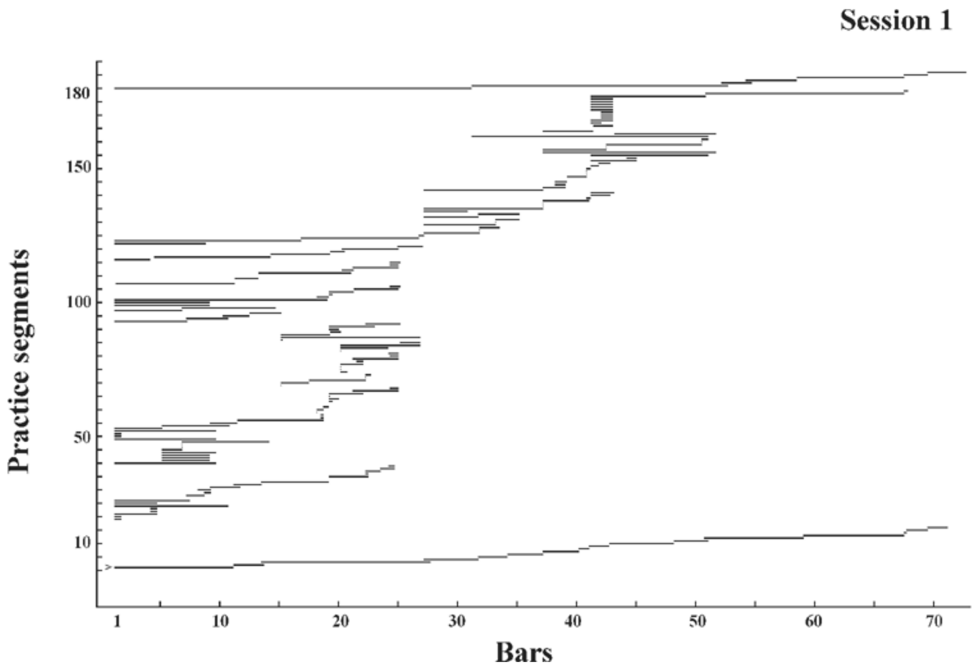


Figure 10.2 Practice record for concert pianist Gabriela Imreh's first practice session of *Clair de Lune* by Debussy. This graph should be read *from bottom up*. Each line represents a continuous stretch of playing through the music. Every time the pianist stopped and picked up again, this is shown as a new line slightly above the previous line.

Source: Chaffin (2007, Figure 3, p. 384), reprinted by permission of University of California Press.

The graph of Gabriela Imreh's practice activity incorporates aspects of 'whole' and 'part' strategies, and shows that she segmented the music into parts corresponding to the musical structure of the work. Advanced students and experts tend to be more intentional about how they segment the music, often dividing a musical work into meaningful structural units such as themes, phrases, sections, and bridge passages (Chaffin et al., 2010; Williamon & Valentine, 2000; cf. Hallam et al., 2012). They also tend to learn the units in a meaningful order, such as principal themes first and then variations, not necessarily in chronological order. By contrast, younger students often segment the composition at points of convenience – such as learning the music line by line (e.g., Hallam, 1997).

When it comes to learning vocal repertoire, a different sort of 'part versus whole' question might arise. Is it more effective to first learn text and music separately, or to learn them together from the start? In previous studies, participants often reported that they learned music and words separately, but these studies relied on data gathered from interviews. Ginsborg (2002) undertook an observational study of 13 classically trained singers in order to find out how singers actually practice. The singers were observed over the course of four to six practice sessions of 15 minutes each as they learned an unfamiliar song. The findings showed that fast, accurate memorizers spent more time practicing words and music *together* during the early (first two) sessions of their practice than slow, inaccurate memorizers. Also, fast, accurate memorizers spent more time practicing music and words *separately* during the final two to four practice sessions than did the slow, inaccurate memorizers.

In sum, the most effective technique was to learn words and music together initially, and then study the parts separately. A subsequent study showed that memorizing words and melody together is an effective strategy for accurate and fluent recall of songs, but only for more advanced singers (Ginsborg & Sloboda, 2007).

Silent practice techniques

Becoming a musician does not simply require developing muscular strength and dexterity. Acquiring expertise also entails cognitive activity, and as such practice does not always need to involve movements. ***Mental practice*** refers to ‘the imaginary rehearsal of a physical skill in the absence of gross muscular movements’ (Theiler & Lippman, 1995, p. 329), and is widely used by athletes in various sports. In music training, it refers to practice that takes place in a person’s mind while imagining the sounds and motor actions involved in playing a musical work.

Mental practice techniques rely to some degree on ***mental imagery***, or the conscious simulation of actual experiences. Imagery can be visual (which one may instinctually associate with the word ‘image’), but can also be auditory or motor. ***Auditory imagery*** involves imagining the sound of music in your head (an example includes ‘earworms,’ discussed in chapter 5), whereas ***motor imagery*** involves imagining the performance of an action or sequence of actions. Neuroimaging research using PET and fMRI suggests that engaging in auditory imagery activates several brain areas associated with both perceiving and performing music, thus providing a neural basis for the potential effectiveness of mental practice (Zatorre & Halpern, 2005).

Analytical pre-study

Two main forms of mental practice techniques have been described: ‘analytical pre-study’ and ‘mental rehearsal.’ ***Analytical pre-study*** refers to mental practice that involves studying the musical score *before* engaging in physical practice (Rubin-Rabson, 1941a). This allows individuals to preview the musical work and analyze its form, identify patterns, and ‘hear’ the music in their heads – in advance of physically playing the piece. Engaging in analytical pre-study prior to physical practice may facilitate smoother, more accurate, and less mechanical playing when sight-reading or playing through a piece for the first time. These advantages likely reflect the fact that the individual builds a mental representation (or performance plan) for the piece as a whole that functions as a mental schema (cf. chapter 5) to guide performance. In particular, pre-study can facilitate rhythmic accuracy when playing complex pieces because the student has time to work out the rhythms first (Rosenthal, 1984). Any technique that can improve rhythmic accuracy facilitates performance, as the most common errors in sight-reading are related to rhythm (McPherson, 1994).

In a study focusing on physical versus mental practice, Bernardi and colleagues asked 15 advanced pianists to learn two excerpts of Scarlatti sonatas, using physical practice to learn one excerpt, and analytical pre-study to learn the other (in a manner that was counterbalanced between participants and pieces) (Bernardi, Schories, Jabusch, Colombo, & Altenmüller, 2013). The pianists were given 30 minutes to study the first piece using the assigned form of practice, and then performed the piece on a MIDI piano by memory. They were then given 10 more minutes to practice the piece – during which time those in the ‘analytical pre-study’ group could engage in physical practice, and those in the physical practice group were asked to continue physical practice only (serving as a baseline comparison). Five days later, the pianists returned and repeated the same procedure with the other piece, using the other form of practice. Their performance was then judged on note accuracy, articulation, phrasing, dynamics, and expressivity.

Bernardi and colleagues (2013) found that the ‘analytical pre-study’ group achieved a performance proficiency of only 40–60% of that achieved by the ‘physical practice’ group, after 30 minutes of pre-study alone. However, when the 30 minutes of ‘pre-study’ were combined with the 10 minutes that allowed physical practice, the results were comparable to those of the ‘physical practice’ group. The researchers conjectured that mental practice may lead to gains that are initially latent, but are quickly manifested after engaging in some physical practice. This is good news for performers who must rehearse when instruments are not available, and for those who are prescribed temporary physical rest while recuperating from playing-related injuries stemming from long hours of practice.

Mental rehearsal

Another form of silent rehearsal is implemented *after* engaging in a period of physical practice, referred to as ***mental rehearsal***. This strategy employs mental imagery after the performer already knows how the piece ‘feels under the fingers.’ Mental rehearsal involves playing through a musical work in one’s mind, using one’s direct experience with the piece to imagine each movement and sound as vividly as possible. It may also involve visualizing parts of the written musical score as if ‘reading’ from it, along with the sounds and motor actions of playing.

Is it beneficial to incorporate some ‘shadowing’ in mental practice – such as moving fingers lightly on the surface of piano keys, or moving the slide soundlessly on a trombone? Bernardi and colleagues identified two main styles of mental practice among advanced musicians: a ‘physical’ style in which the musician often makes fine movements soundlessly (e.g., moving fingers or lips silently or gesturing) to support the formation of their mental representations of the piece, and a ‘mental/analytic’ style in which the musician does not make many overt movements and focuses on a formal analysis of the musical score (Bernardi et al., 2013). Neither style seemed to be better than the other for producing a more proficient performance. Abilities such as how effectively the musician engages in auditory imagery (cf. Highben & Palmer, 2004) seem to play a more critical role in mediating the effects of mental practice. In other words, the ability to vividly imagine the music playing in one’s mind was more important to the effectiveness of mental practice than whether or not the musicians incorporated some (soundless) physical movements.

Overall, studies suggest that physical practice seems to be a more effective means of improving performance than mental practice alone, especially at the early stages of acquiring musical skills (e.g., Lim & Lippman, 1991; Pascual-Leone et al., 1995). However, a *combination* of physical practice and mental practice may be equally as effective for improving performance as physical practice alone (e.g., Bernardi et al., 2013; Coffman, 1990). For instance, Ross (1985) showed that improvement in pitch, rhythm, and articulation was similar for slide trombone players using a combination of physical and mental practice and physical practice alone. Further, interspersing physical with mental practice in one’s routine was found to be superior to pure physical practice for facilitating stronger memorization of keyboard music in advanced pianists (Rubin-Rabson, 1941b) and for developing tonal quality in singers and pitch accuracy in guitar players (Theiler & Lippman, 1995).

It should be noted that most studies have employed only short-term exposure to mental practice. Mental practice itself takes practice, and further study is needed to determine the effects of long-term engagement in pre-study and mental practice.

Aural and audiovisual models

Modeling is a teaching tool in which a specific behavior is demonstrated and is then imitated by another person. Until the invention of recording technology, music students were dependent on live models to demonstrate a skill for a solo student or ensemble to emulate. Today, teachers can easily equip students with ***aural*** or ***audiovisual models*** in the form of audio or video recordings to use during music practice. Rosenthal (1984) found that advanced instrumentalists who listened attentively to a recording of a musical work before physically practicing it for three minutes outperformed those who had only physically practiced the piece for 10 minutes. Specifically, those who listened to the aural model first received significantly higher scores for correct notes, correct rhythms, accuracy of tempo, and expressive playing (dynamics) on their subsequent performance of the piece than those who had practiced physically for more than three times as long. Singers, in particular, benefit greatly from mental practice combined with an aural model. Theiler and Lippman (1995) found that interspersing physical practice with listening to an aural model significantly improved pitch accuracy, dynamics, tempo, and tonal quality of vocalists' performance of a new piece (compared to physical practice alone, or to physical and mental practice without an aural model).

The model does not have to be instrument-specific to be effective. Rosenthal's (1984) study used an audio recording of a violin performance as the aural model for brass and woodwind instrumentalists. Most students seem to be able to transfer the skills to their respective instruments (Dickey, 1991), perhaps with the exception of beginning students (Linklater, 1997) and young children (Green, 1990). Young children, in particular, may need to follow a model that is as similar as possible to their own voices or instruments. For instance, Green found that first-grade through sixth-grade elementary school children were most accurate in imitating sung pitches when following a child vocal model, less accurate when imitating a female adult vocal model, and performed least accurately when the vocal model was an adult male.

Not all studies show that practicing with aural or audiovisual models significantly improves performance (e.g., see Morrison, Montemayor, & Wiltshire, 2004), as the effects of modeling are likely to be mediated by individual, developmental, and social factors. For example, Linklater (1997) exposed beginning clarinet players to one of three modeling conditions during eight weeks of practice: (1) a videotape of a clarinet performance, (2) an audiotape of the same clarinet performance, or (3) an audiotape of other instruments playing an accompaniment (as a control). After eight weeks, students in the clarinet videotape model group (using both audio and visual modeling) scored significantly higher than the other groups when judged on several aspects of technique (embouchure, hand position, instrument position, and posture). However, there were no significant differences between the groups with respect to the students' performance of the music. Further, Linklater conjectured that the improvement in the videotape group may have been mediated by greater parental involvement, as parents (who often did not play clarinet themselves) may have felt more comfortable reminding their children about proper technique with the support of an audiovisual aid.

Studies have shown that music teachers at elementary through conservatory levels often encourage students to listen to recordings of masterful performances, but rarely integrate audio or audiovisual recordings of music into private lessons (Gaunt, 2008). This is rather surprising, as the widespread availability of smartphones, tablets, and Internet resources (such as YouTube) puts a range of aural and audiovisual models at the fingertips of many teachers and students. Also, a student's smartphone can easily be used to record the teacher modeling a technique during a private lesson or ensemble practice, for the student to review when practicing at home. In addition, this recording may equip parents with an audiovisual model to reinforce the teacher's instructions at home (as noted in the discussion of Linklater's 1997 study). It may be up to this generation of 'digital natives' (Prensky, 2001) to acquaint their teachers with the rich technological advances that may facilitate the use of aural and audiovisual models in private and ensemble lessons.

The risks of intensive practice: Musicians' injuries

Considering the *thousands of hours* of practice required to become proficient at music, it is not surprising that a large number of amateur and professional musicians report pain and other ailments related to playing their instruments. The reported prevalence of pain and other conditions varies (depending on the duration and severity of the symptoms), but reports are as high as 86% of orchestral musicians for 'pain,' and 41% for 'disabling pain' experienced within the last 12 months (Leaver, Harris, & Palmer, 2011). A broad category of conditions affecting musicians is referred to as ***playing-related musculoskeletal disorders (PRMDs)***. These encompass a wide variety of conditions affecting the muscle and bone, including: *repetitive strain injuries*, *muscle-tendon overuse syndromes* (e.g., tendinitis), *nerve entrapments* (e.g., carpal tunnel syndrome), *muscle imbalance* (caused by overtraining some muscles in relation to opposing muscles, leading to pain and limited movement), and *focal dystonia* (a rare neurological condition, discussed in the final section of this chapter, associated with loss of control of specific muscles). Other conditions frequently reported by musicians are discussed elsewhere in this book (including *music-induced hearing loss* in chapter 3, and *music performance anxiety* in chapter 12).

The incidence of PRMDs is higher in females, in individuals with smaller frames, and rises sharply after age 30 years (Zaza, 1993). Lifestyle and habits, such as poor nutrition or substance abuse, and bad posture and poor playing technique, and personality traits such as high trait anxiety and introversion, are also associated with higher incidence of PRMDs (Zaza, 1993). In singers, smoking and drinking caffeine or alcohol tends to dry out the vocal cords, which makes them more prone to strain and injury, increasing the risk for singers to develop nodules (callous-like growths on the vocal cords, formed out of scar tissue as injuries heal). Repertoire requiring musicians to strain while performing, such as hyperextension of the pianist's hand-span while playing wide intervals for extended passages, or straining the voice to sing beyond one's comfortable range for long periods, also increases risk of injury (Quarrier, 1993). The fear that injuries will be perceived as a sign of weakness or poor musicianship leads many orchestral musicians to conceal their injuries and delay seeking treatment (Rickert, Barrett, & Ackermann, 2014), so psychosocial factors also contribute to the development of chronic and debilitating conditions.

Playing instruments that necessitate ***asymmetrical*** postures in order to play (such as flute, violin, and bassoon) also increase the risk for developing PRMDs compared to instruments involving more ***symmetrical*** postures (such as piano, clarinet, and trumpet). Asymmetrical postures distribute body weight unequally between the left and right sides of the body, and involve more curvature of the spine. Accordingly, pain and discomfort in the neck, shoulders, and upper back are especially prevalent in flute players (Lonsdale, Laakso, & Tomlinson, 2014) and in violinists and viola players (e.g., Lahme, Eibl, & Reichl, 2014). These instrumentalists spend prolonged periods of time holding both arms up and extended to one side of the body. Further, a study of 10- to 18-year-old students at a music conservatory in Italy showed that students who play an asymmetrical instrument are more likely to adopt a non-optimal posture while playing compared to those playing symmetrical instruments (Ramella, Fronte, & Converti, 2014). In turn, playing an instrument with non-optimal posture significantly increases one's risk for developing musculoskeletal pathologies.

Playing instruments while seated (as is typical in ensembles and orchestras) can compound strain on the body, especially for those playing asymmetrical instruments. Solo instrumentalists who play asymmetrical instruments naturally compensate for the imbalanced distribution of weight by making adjustments to the pelvis when performing while standing, but these movements are restricted when playing while seated. One's position in relation to a shared music stand may also contribute to the risk for developing PRMDs. A study using 3D motion capture analysis and a platform to measure weight distribution showed that the violinist sitting to the right of a shared music stand bears a significantly more imbalanced load on the spine, and a diminished range of motion in the right (bowing) arm (Spahn, Wasmer, Eickhoff, & Nusseck, 2014). In comparison, violinists sitting to the left of a shared music stand tend to adopt a more balanced sitting position as they lean to the right to view the music, which compensates for the violin resting on the left shoulder and positioning of both arms to the left side of the body. Crowded ensemble seating also increases tension, constraining movements of the extended arms of instrumentalists such as flute players and violinists (Lonsdale et al., 2014; Spahn et al., 2014).

Treatment and prevention

Musicians' injuries are so widespread and can be so debilitating to one's career that a specialty in Performing Arts Medicine has been established (focusing on injuries in musicians, dancers, and other artists), along with a professional society known as the Performing Arts Medicine Association and its journal, *Medical Problems of Performing Artists*. Centers for the treatment of medical problems specific to musicians, dancers, and other artists typically rely on a team-based approach; clients are treated by a multidisciplinary team of physical therapists, occupational therapists, psychologists, neurologists, and other specialists. Treatment programs based on principles yielded by research studies, such as evidence-based physical therapy management programs (e.g., Chan & Ackermann, 2014) are also taking hold.

Good technique, correct posture, proper diet and hydration, and frequent rest breaks serve as protective factors against injury (Zaza, 1993). In particular, good technique does not just constitute good form, but facilitates optimal posture for balanced distribution of weight, mobility, and breath support. Poor technique, on the other hand, can cause greater wear and tear over time due to inefficient action and strain. For example, in flute players, tilting the head, jutting the chin either too far forward or downward, or hunching the shoulders puts greater strain on muscles and joints while playing, and diminishes lung capacity and diaphragmatic support (Lonsdale et al., 2014). Poor posture in singers greatly affects breathing and increases strain on the voice, and the tendency to move the head forward when singing higher pitches increases stress on the muscular, spinal, and neurovascular systems (Quarrier, 1993). Stretching exercises (full body stretches, warm-up vocal exercises, and scales) also lower the risk for developing PRMDs. One study showed that only 36% of string players who routinely stretch before engaging in practice reported PRMDs, as opposed to 69% among those who do not regularly stretch before practice (Kim, Kim, Min, Cho, & Choi, 2012).

Musicians subject their bodies to intense physical challenges, as do athletes. They often underestimate the physical and psychological toll exacted by the thousands of hours spent performing repetitive motions, holding the same posture for extended periods, and sustaining the force required to play their instruments. Healthy habits must be developed along with musical skills, to ensure the effectiveness of deliberate practice and the longevity of a musician's career.

Music practice and the brain

In the final section of our chapter, we turn our focus to neuroscientific studies on the effects of intensive music practice. Music practice is of interest to neuroscientists because of its implications for ***neural plasticity***, which refers to the impressive capacity of our nervous system for change (as discussed in chapter 4). The idea that intensive motor training can induce *functional* changes in the brain (i.e., modification of the organization of neuronal networks to increase efficiency) has been quite well documented (e.g., Pantev, Engelien, Candia, & Elbert, 2003; Dalla Bella, 2016). Other studies have found evidence that increased use can also lead to *structural* changes in the brain (Pascual-Leone, 2003). In other words, training may not just bring about change on the level of neuronal reorganization, but may also guide the way the brain develops. In the final section of this chapter, we consider whether musical practice brings about structural and functional change in the brain independently of one's genetic make-up.

When considering plasticity, it is important to address whether musicians are born with structural features that predispose them to musical ability, or whether intense training over a long period of time can induce structural adaptations in the brain. Although research in this area is relatively new, there is accumulating evidence that structural changes are due to 'use-dependent' changes rather than differences present at birth. For instance, Gaser and Schlaug (2003) found a significant correlation between level of music training (namely, professional musician, amateur musician, and nonmusician) and amount of increase in gray matter volume in the motor (precentral gyrus), auditory (Heschl's gyrus), and visual-spatial (parietal) regions. Specifically, professional pianists had the highest gray matter volume in these regions, and nonmusicians had the lowest volume. (Further discussion of gray and white matter can be found in chapter 11 in this volume, in the section on music performance and the brain.) The researchers attributed their finding of higher gray matter volume in professional musicians to 'structural adaptations in response to long-term skill acquisition and the repetitive rehearsal of those skills' (Gaser & Schlaug, 2003, p. 9240), rather than to innate differences.

Hyde and colleagues have also undertaken a longitudinal study focusing on a group of 6-year-old children who received private keyboard training. Their development was compared to another group (matched as closely as possible in gender, age, and socioeconomic status) who received a weekly music class in school, but did not receive instrumental training (Hyde et al., 2009). Baseline measures at the start of the study showed no pre-existing neural or cognitive differences between groups. However, after 15 months, children who took keyboard lessons performed significantly better than controls on tests of fine motor skills and auditory discrimination skills. Moreover, the authors observed increased neural development in brain areas associated with keyboard performance, including the hand area of the primary motor cortex and the corpus callosum. These findings help confirm that results from earlier cross-sectional studies (i.e., those comparing age groups that comprise different sets of individuals) truly reflect changes that are caused by practice.

The findings of a few studies suggest that some structural adaptations may be instrument-specific. For example, Elbert and colleagues found that the region of the somatosensory cortex that represents input from the left hand was significantly more responsive to tactile stimulation in a group of expert string instrumentalists than in nonmusicians (Elbert, Pantev, Wienbruch, Rockstroh, & Taub, 1995). This may be because the fingers of the left hand may be more dexterous among professional string players due to training. Further, the effect was more pronounced for the string instrumentalists' left pinky finger than for the left thumb. This may be due to the fact that the thumb is used mainly to hold the neck of the instrument, whereas the fingers (including the pinky) are constantly in motion (as shown in Figure 10.3).

Figure 10.3 Neuroscientists have found evidence of increased responsiveness in the region of the somatosensory cortex that represents input for the left hand (especially the pinky finger) in expert violinists after years of playing an instrument which requires dextrous movements of the fingers of the left hand, and particularly the little finger.

Source: Photograph by John Lacko, used with permission of Barry Ross and the Kalamazoo Symphony Orchestra.

Elbert and colleagues attributed the increased responsiveness of the activated area to the possibility that the brain assigns more neural tissue to the processing of the hand and fingers that are more heavily used. In contrast to players of stringed instruments, pianists may favor the right hand. Bangert and Schlaug (2006) investigated the shape of instrumentalists' precentral gyrus, the fold that separates the motor cortex from the parietal lobe (see Figure 4.2). The shape of this fold suggested greater density of the motor area in the left hemisphere (which controls the right hand) for pianists, and greater density of the motor area in the right hemisphere for violinists.

Further, the most pronounced differences have been found for musicians who commenced training at a young age, most likely because intense training took place during a time when the brain was particularly malleable. In a study by Schlaug and colleagues (Schlaug, Jäncke, Huang, Staiger, & Steinmetz, 1995), the size of the anterior corpus callosum was larger in a group of professional musicians than in a group of matched controls, but this difference was larger for a subgroup of their musicians who had begun music training at 6 years or younger. The development of this anatomical change has been confirmed in subsequent research that involved measuring structural changes in a group of children before and after 15 months of musical training (Schlaug, 2009). However, although effects seem to be greater for musicians who started music training early in life, differences are sometimes found for 'late starters.' In Elbert and colleagues' (1995) study, string players who had not started music study until after age 12 still showed higher levels of electrical activity in areas corresponding to the left pinky finger compared to nonmusicians. As with so many areas of human development, it appears that there is a more flexible 'sensitive' (rather than rigid 'critical') period for the brain to adapt to intense motor training.

Effects of practice on music perception

Music training can also shape how your brain responds to incoming sounds, even when not performing. The findings in this area are remarkable, in that changes in neural responses have been found after only 20 minutes of practice among individuals with no prior musical training. In one study, participants were presented with melodies they had just physically practiced as well as melodies that resulted from simply re-arranging the order of pitches in previously learned melodies (Lahav, Saltzman, & Schlaug, 2007). Neural responses were greater, specifically in Broca's area, when listening to previously learned melodies that the participants had physically practiced. (It should be noted that Broca's area was once thought to be a 'language area,' though it is now believed to serve multiple purposes.) In another study, only 20 to 30 minutes of physical practice elicited larger neural responses (in similar areas) to melodies that had been practiced versus melodies that had been learned merely through listening without playing (Mutschler et al., 2007). These findings suggest that performers build associations between actions (key presses) and their effects (pitches) rapidly, thus providing a sensorimotor basis on which to build further musical expertise.

A more recent study by Ellis and colleagues (2012) examined neural responses to pitch and rhythm changes in both children and adults, and used statistical analyses to separate the influence of musical training from the effects of age. The authors found separate brain areas that changed as a function of age versus experience. In particular, musical training enhanced responses in the rear of the superior temporal gyrus.

These rapid associations early in practice may be based on what has been dubbed the brain's 'mirror neuron' system. This appears to be because musical training causes performers to engage in mental simulation while listening. In the 1990s, researchers discovered (by accident) that monkeys' brains possess cells that have since been dubbed ***mirror neurons*** (Rizzolatti & Craighero, 2004). Specifically, while conducting research focusing on neurons that regulate motor behavior (in this case, reaching for a food pellet), Rizzolatti and Craighero found certain cells that responded similarly regardless of whether the monkey reached for the pellet or merely observed an experimenter reaching for the pellet. It appears as though a similar system operates in humans for some activities, including music learning, although this claim has been met with vigorous debate and must be regarded as tentative for now (e.g., Gallese, Gernsbacher, Heyes, Hickok, & Iacoboni, 2011).

Musicians' focal dystonia

As we have seen, practicing music can have profound effects on the brain's structure and function. But there can be a darker side to practice that is too intensive, particularly when it is massed rather than distributed. Earlier we discussed the problem of peripheral injury that can occur if one practices too fervently in a short period of time, leading to the risk for developing PRMDs. Although such conditions can be highly disruptive to one's performance career, possibly the most debilitating kind of injury has a neural origin: ***musician's focal dystonia*** (Altenmüller, Ioannou, & Lee, 2015). At first appearance, musician's dystonia may seem to be a very severe kind of cramp. Pianists who suffer may find that their fingers curl up once they start to play, and become nearly unusable. The embouchure of a trumpeter may be affected by uncontrollable spasms. However, despite the peripheral manifestation of musician's dystonia, it does not seem to arise from muscular or nerve problems at the periphery. Its origin seems to be in the brain. As a result, musician's dystonia is tremendously hard to reverse, leading some experts to recommend that students who suffer from musician's dystonia seek an alternate profession (Jabusch & Altenmüller, 2006a).

The cause of musician's dystonia is of interest because it reveals the potential hazards of practice. Practice can lead to the development of two opposing forces in the brain. As we discussed earlier, practice enhances the neural representation of muscles used to perform music, such as the area of the primary motor cortex used to control finger movements. However, the brain must also be able to inhibit movements of muscles that are *not* used for individual movements. So when your brain commands the index finger to move toward a keyboard, it must make sure that the signal does not also lead to contractions of the middle, ring and pinky fingers.

It seems that musician's dystonia happens when both sides of the overall mechanism do not develop, and the brain has excitation without inhibition (Rosenkranz et al., 2005). This is particularly likely when one practices an instrument intensively over a short period of time, as in massed deliberate practice. In fact, Altenmüller (2005) suggests that Robert Schumann's famously aborted career as a concert pianist may have been due to musician's dystonia, as opposed to other peripheral causes that have been proposed (e.g., stretching his ring finger while sleeping by tying a string around it that hung from the ceiling, perhaps in an attempt to improve the flexibility of this notoriously inflexible finger). More recent documented cases of focal dystonia include pianists Gary Graffman and Leon Fleischer and oboist Alex Klein. Fortunately, musician's dystonia is an exceedingly rare condition; the estimated incidence is only about 1% in professional musicians (Jabusch & Altenmüller, 2006b).

Coda

As a whole, the literature on the acquisition of musical expertise gives us reason to question the popular notion that 'some people are just born with it.' At the same time, the role that talent and innate characteristics may play in the pursuit of excellence in any field cannot be overlooked. Increasingly, theorists and researchers are seeking to understand how important components such as genetics, deliberate practice, personality, and other individual factors, social support, and many other variables interact in complex ways to produce individual outcomes in musical achievement.

Note

1 The first author of this article is Yi Ting Tan, no relation to the co-author of this book.

11 The psychology of music performance

Glenn Gould changed how we think about Bach, twice. The eccentric Canadian pianist, known for singing along with the melody in his performances (which you can hear in his recordings), recorded Bach's *Goldberg Variations* when he was 22, and another time just before his death in 1981. Bach is typically performed with precise timing that closely reflects notated durations and is played with a fairly consistent loudness (intensity) across notes. Gould adopted an entirely different strategy in his first recording by performing the variations with the kind of expressive timing and fluctuations in intensity that one usually associates with Romantic composers like Beethoven or Chopin. His radical interpretation breathed new life into a previously obscure Bach piece, and his recording remains incredibly popular to this day. Gould's later 1981 recording is a fascinating reaction against the young man's formerly more Romantic interpretation. The 1981 recording more closely conforms to the more restrained standard Baroque practices, in many cases exaggerating the temporal precision of his performances. Taken together, these two recordings represent the extent to which a performer's unique interpretation(s) can influence how we hear music.

Note that the reinterpretation of Bach that is present in Gould's recordings had nothing to do with changing the basic content of the music. Every note remained as in the original scores, and each notated duration was represented similarly in the performance (e.g., eighth notes were not played as if half notes). What Gould, like other performers, did was to vary how the basic content is expressed. This is not unlike how we produce language through speech: Consider how one might say 'wasn't that a terrific lecture' with different implied meanings. Note also that the use of expression in music performance is not confined to classical music. Consider Eric Clapton's acoustic rendition of 'Layla' from the 1990s: Originally an upbeat rock song by Derek and the Dominoes, the song was reinterpreted as a soulful blues tune. Similarly, the Lennon–McCartney classic rock song 'We Can Work It Out' takes on radically different funk vibe in Steve Wonder's interpretation.

More basic questions about performance are worth considering, too. How does the performer put it all together in his or her head? Songs and longer pieces include hundreds or even thousands of notes to remember. Some rock bands perform three-hour concerts without consulting a single sheet of music notation or chart of guitar chords. Skilled pianists generate finger movements separated by time intervals that are as small as 10 milliseconds, all while timing the notes with great precision. Moreover, performers need not be professionals, and they need not be pleasing to hear! We all make music now and then, even if that means singing in the shower.

The focus of the current chapter is on the cognitive and motoric processes that contribute to an individual's performance of music. Other aspects of performance – such as the influence of the performers and audience members on each other, the social dynamics of ensemble and group interactions, and the topic of music performance anxiety – are discussed in chapter 12. We trace the progress of a performance in three stages: the formation of a performance plan by using short- and long-term memory, controlling motor actions involved in performance, and monitoring the outcome of one's actions. We follow this discussion with a brief summary of cognitive neuroscience research on performance.

The role of memory in music performance

When you think of the role of memory in music performance, what probably springs to mind is the considerable challenge of retrieving a long piece of music from memory without any retrieval aids, such as music notation. This is indeed a considerable feat, but it is not the only way in which we use memory in a performance. In fact, there are memory demands (albeit of a different sort) even when we play while reading notation. In order to illustrate the different kinds of memory performers use (drawing on many decades of research in cognitive psychology), we consider three performers: Margaret, who plays a long piano sonata from memory; Nancy, who is sight-reading from music notation; and John, who is improvising in a jazz ensemble.

Long-term memory

Long-term memory, as the name implies, holds information over a long time-span, which can be minutes, days, or one's entire life. Its capacity may be limitless; at least no one has identified a limit to how much information long-term memory can hold. This kind of memory is most apparent in Margaret's performance, though as we will see, there are elements of long-term memory in Nancy and John's engagement with music as well. When Margaret learns her piano sonata and plays it without notation (a practice initiated by Franz Liszt in the 19th century), she relies on long-term memory.

You may be surprised by the claim that long-term memory has infinite capacity. We have all had the experience of retrieval failure, and the sense that some memories may have disappeared over time. Memories are not like files in a file drawer. They are usually not stored in a way that precisely reflects the original physical information. Instead, they may be stored using cognitive shortcuts that are subject to limitations, which is why eyewitness testimony can be unreliable. One shortcut that music performers use is called ***chunking***. Chunking involves grouping together information into larger, more meaningful units. For instance, the number sequence 177614921945 can easily be remembered if it is grouped in the following chunks: 1776 (year the declaration of independence was signed), 1492 (when Columbus came to America), and 1945 (the end of World War II). Similarly, musical phrases typically form chunks in memory (Miklaszewski, 1989; Williamon & Valentine, 2002). In fact, effective musical practice involves segmenting a long musical piece into phrase-based chunks, which we discussed in chapter 10.

When Margaret retrieves the notes and chords from her sonata during her memorized performance, she may rely on two kinds of long-term memories. The most likely kind of memory she uses is referred to as procedural memory. ***Procedural memory*** involves memories for motor skills, sometimes called 'muscle memory.' Procedural memories do not usually involve conscious awareness. An example related to piano performance is touch-typing. If you are able to touch-type, try to recall what finger is used to press the 'e' key. It is likely that this task was more difficult for you than typing the word 'the,' and, more to the point, it is likely that you recalled which finger you used by imagining your fingers striking the keys. Similarly, pianists are often not consciously aware of which finger and what key they are pressing during a fluid performance.

There is a drawback to relying entirely on procedural memory. The retrieval of procedural memories is highly contingent on serial order, and relies on a process described as ***associative chaining***, which means that retrieval of a note or chord at the present time is based on associations with immediately preceding notes and chords (Chaffin, Demos, & Logan, 2016). The problem occurs when one link in the chain breaks. One of your authors (PQP) experienced this at a piano recital. In the middle of one of the movements of Beethoven's first piano sonata, memory for what came next was suddenly lost, and it took several minutes to retrieve this memory and continue the performance.

Thus, it is important also to have access to some kind of explicit, or consciously accessible, long-term memory. Chaffin and colleagues refer to this kind of memory as ***content addressable***, because it allows a performer to access specific content from any point in a musical piece. For example, can you recall the very last phrase of 'Happy Birthday' without singing the whole song in your mind? A famous feat of content addressable memory comes from the conductor Arturo Toscanini (Marek, 1975). He was reported to have reassured a bassoonist with a broken key on his instrument that the key in question would not be used during that evening's concert, having drawn not only from his precise memory of the music for that concert, but also from the physical actions necessary for that specific instrument!

Let us now consider the improvised performance of John. At first glance, improvised performances may seem to be completely spontaneous. Although we know much less about improvisation than we do about performance from notation (Ashley, 2016), what we know so far suggests that improvisation does rely on memory. First, most practicing jazz musicians will dispute the notion that jazz solos are entirely spontaneous, based on the simple fact that jazz musicians *practice* certain elements of their solos (Berliner, 1994). Specifically, soloists often practice short motifs that can be combined in a vast number of ways, a claim that was verified in analyses of solos by Charlie Parker (Norgaard, 2014).

In this way, jazz solos mirror language. Although most statements are creatively novel when taken as a whole, they typically consist of shorter phrases that we use quite often. You may never have said 'Take the garbage and throw it off the George Washington bridge' in your life, yet it is possible that you have produced all the constituent words before and perhaps even short phrases from it (e.g., 'Take the garbage'). A second way in which jazz solos involve memory stems from the fact that they incorporate notes that abide by a certain rule system, referred to as a musical grammar, that fits within the genre. Jazz musicians thus rely on a ***schema*** (Johnson-Laird, 2002), or rule-based information in long-term memory to guide their improvisations during performance (see chapter 5 for more discussion of schemas).

In keeping with this idea, a recent study of musical improvisation on the piano showed that the notes improvisers chose are constrained by the key in which they improvised, thus supporting the role of tonal schemas (Goldman, 2013). Interestingly, this tendency was magnified when pianists improvised in a key that was less familiar and more demanding with respect to fingering (the key of B as opposed to B♭ – the former includes more 'black keys' on the piano, which makes performance more technically demanding). Solos were also more repetitive when played in an unfamiliar key, a tendency that is further magnified when soloists are asked to divide their attention between the task of improvising and some other cognitive task (Norgaard, Emerson, Dawn, & Fidlon, 2016). Thus, performances may be more predictable (closer to the schema) when the context is more challenging for a performer.

What about Nancy, who is playing an unfamiliar piece by reading notes directly from music notation? She can rely entirely on a visual representation of the piece (the notation) and possibly eliminate any need for long-term memory. Although this is possible in principle, in practice, good sight-readers do not rely entirely on visual notation. In order to sight-read skillfully, one must read notation at an incredibly rapid rate, and often this process is guided again by schematic long-term knowledge to help the reader anticipate future content. The fact that this occurs is best demonstrated by cases in which using prior information causes the performer to misinterpret actual content. One example is a phenomenon known as ‘proofreader’s error,’ a term borrowed from the typing literature. The task of proofreading is paradoxically made more difficult by being a good reader. A good reader uses context based on sentence structure and text content to make predictions for upcoming words, and therefore increase reading speed. By doing this, however, the reader may skip over certain words or even mistakenly perceive misspelled words as being correctly spelled, leading to errors of typing.

The same thing happens in skilled sight-reading for music. Sloboda (1985) recounts a real-life anecdote in which a misprint in the notation for a piece by Brahms went unnoticed for years until a *poor* reader performed it *exactly as notated* during a piano lesson. After some discussion, the teacher inspected the notation to see that in fact the student was reading correctly, and the notation was in error. In an earlier paper, Sloboda (1976) demonstrated similar errors in sight-reading for obscure classical pieces in which selected notes were altered to be incorrect in an obvious way. Pianists incorrectly ‘corrected’ these notated alterations about one third of the time, and in fact made more errors from this subclass on a second sight-reading trial in which their overall performance improved!

Although reading from notation does use long-term memory to some extent, the greatest demand placed on memory in this task is not of the long-term sort. The main challenge in reading notation fluently is the ability to read in advance and hold onto that information for a short period of time as one plays. This leads us to a different form of memory that performers use: working memory.

Working memory and planning of serial order

Working memory involves the temporary storage of consciously accessible information (cf. Baddeley, 1986; Miller, 1956). It is used when our minds sustain multiple tasks simultaneously, such as when a pianist must plan a rapid sequence of keystrokes. Working memory allows the performer to retain all these keystrokes at once, which allows her to generate actions for performing each note rapidly. If a performer were to retain only one note at a time (thus minimizing any role of working memory), the resulting performance would likely be slow, halting, and displeasing. Only a beginner would perform in this way. Better performers will use working memory to enhance the fluidity of performance.

How much information do skilled readers store in working memory? Some early evidence comes from further studies of music reading by Sloboda (1974), who adapted a procedure from research on reading. One way to determine how much information a pianist has in working memory is to remove the notation at an unpredictable point while they are performing, and then observe how much longer production can continue (a similar procedure in reading involves removing the text as one reads aloud). The number of notes that can be performed after the notation is removed constitutes the ***eye–hand span***, and is a measure of working memory capacity for sight-read material. Skilled sight-readers were found to have an eye–hand span of 10–11 notes, whereas the span for less skilled sight-readers was 4–6 notes, which is closer in line with estimates of basic working memory capacity (commonly considered to be 7, but probably closer to 4 items; Cowan, 2015).

Despite the usefulness of the eye–hand span, it cannot serve as the only measure of working memory in performance. First, sight-reading skills are distinct from other music performance skills (Lehmann & Ericsson, 1996). Second, not all the items in working memory may have the same priority. Certain notes or chords in music may be more salient in memory than others. As such, we turn to what is perhaps a counterintuitive source of evidence regarding the content of working memory and priority assigned to different items: the occasional mistakes performers make.

One of the most influential papers in early cognitive psychology was 'The problem of serial order in behavior' by Karl Lashley (1951). In it, he reflected on the significance of what kind of 'serial ordering errors' people make when they speak. A ***serial ordering error*** happens when you misplace the order of a particular event while maintaining the appropriate content. Here is an example from speech, heard once by one of your authors (PQP), with the critical components underlined:

Spoken:

> 'We are here to celebrate the wedding of Kelley and Sheith . . .'

Intended:

> 'We are here to celebrate the wedding of Shelley and Keith . . .'

In this instance, the priest at a wedding swapped the position of two speech sounds (phonemes): the 'Sh' sound from 'Shelley,' and the 'K' sound from 'Keith.' The priest produced all the correct phonemes, but misplaced the order of two of them.

Errors like this are unlikely to result from simple mispronunciations. If so, then the 'K' of 'Keith' (an unvoiced velar stop) should have been replaced with a phonetically similar sound, like the hard 'G' (a voiced velar stop). Instead, these errors reflect the content of working memory. (Interestingly, Freud also observed that such 'slips of the tongue' probably involve deeper processes, although his explanation drew on different theoretical assumptions!) The fact that the speaker swapped sounds from the beginnings of words is significant. When retrieving the word 'Shelley,' he was also thinking about the next proper noun 'Keith,' and was breaking down each word into its components: syllables, and phonemes. Thus, this error reflected the hierarchical structure of words, and also provided an indication of how far ahead (two words) the content of working memory extended.

Serial ordering errors occur in music performance as well. Research to date has focused on piano performance, and it appears as though 'slips of the finger' do not merely result from 'missing' the intended key and instead hitting the one next to it. Just like errors in speech, musical errors seem to reflect mental associations of content (in this case, the notes) with the global structure of music. The intended note is typically referred to as the ***target***, whereas the note that slips in as an error is called an ***intruder***. Two examples of serial ordering errors are shown in Figure 11.1. The second row in this figure illustrates an ***anticipatory error***. In such cases, an event is performed before its intended location. In the example, the D from measure 3 replaces the E from measure 2. The bottom row shows the opposite sort of error, a ***perseveratory error***. In the example, the A from measure 1 is repeated in measure 2, replacing the intended G. Note that anticipations and perseverations both involve repeating the intruder; it appears both in its intended location and at that target's location. One difference

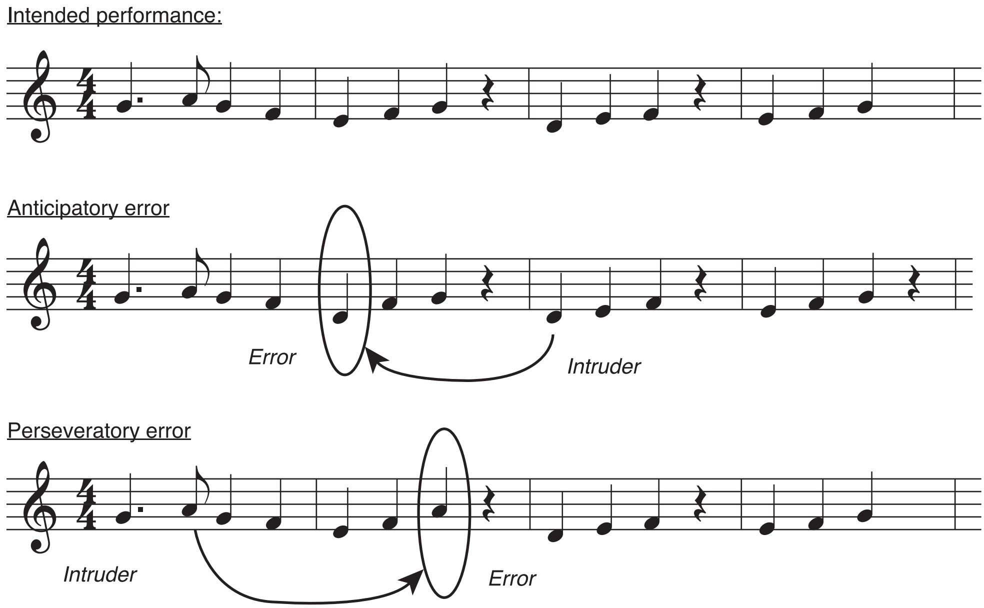


Figure 11.1 Examples of serial ordering errors in the performance of 'London Bridge Is Falling Down' (first phrase).

between speech and music is that musical errors rarely 'trade' places as in the Shelley/Keith example (referred to as an ***exchange error***).

Like speech, serial ordering errors in music reflect its structure. In most piano pieces, the two hands play different parts, and in some 'polyphonic' music, each hand may play a distinct melody line. In such cases, Palmer and van de Sande (1993) found that notes from within the right hand's melody tend not to replace notes from within the left hand's melody. In a subsequent study, Palmer and van de Sande (1995) demonstrated that errors only travel within a single musical phrase. Thus, if one makes an error within the first phrase of 'Twinkle, Twinkle, Little Star,' the intruder would probably not come from one of the notes associated with the second phrase, 'how I wonder what you are.' The implication of both these findings is that a performance plan is divided into voices and melodic phrases, as illustrated in Figure 11.2.

A broader implication of these findings is that although a musician may have the entire piece stored in memory, the process of retrieval during performance does not result in the performer retrieving the entire piece as one large unit. Nor does the performer retrieve the piece one note or one chord at a time. Instead, memory retrieval is 'incremental' in nature. At any given point in time during a performance, the performer holds some, but not all, notes from the piece in working memory. The notion of incrementality, like analyses of serial ordering errors, originated in speech production research as a way of accounting for the fact that people seem to plan what they want to say at the same time as they are actively speaking (e.g., Kempen & Hoenkamp, 1987).

In addition to the structural boundaries shown in Figure 11.2, performers prioritize events within these boundaries differently as they progress through a musical piece. In other words, at any given time, some notes within a phrase are more accessible than others.

The 'Range Model' of planning accounts for incremental use of working memory in piano performance (Palmer & Pfordresher, 2003). This mathematical model predicts the accessibility

Figure 11.2 A schematic illustration of structural boundaries in a performance plan.

Source: Adapted from Palmer and van de Sande (1995, Figure 2, p. 951) with permission of C. Palmer and the American Psychological Association.

of musical events during planning of a musical performance based on two assumptions. First, notes that are closer to the current position are more accessible than notes that are farther away. For instance, if the current note is in position 5, then notes at positions 6 and 4 will be more accessible than notes at positions 8 and 2 (see Figure 11.3, lower left). Second, notes are more accessible if they are associated with a similar metrical accent to other notes (see chapter 6 for more on metrical accents). Thus, when one plans to produce a note or chord on a 'weak' beat, notes and chords on other weak beats are more accessible in working memory than those on strong beats (see Figure 11.3, lower right). The combined predictions of these two components are shown in the top panel in Figure 11.3. As such, intruders are more likely to come from nearby positions that share a similar metrical stress to the current position.

The top panel of Figure 11.3 illustrates what the performer may have on his or her mind during a performance (an elaboration on the thought bubble shown in Figure 11.2). The grid of X's below the music illustrates the pattern of metrical accents at each position, with more X's denoting stronger accents (see chapter 6 for more detail). The current position, highlighted by the rectangle superimposed on the notation, is a strong metrical position. The bars in the graph above the notation illustrate how accessible or 'active' each note is in working memory. In other words, the bars represent the likelihood that each note will be performed at the current position.

In Figure 11.3, the current note, indicated by the bar extending above the plot, is most highly active, and thus is the most likely to be performed. However, other positions are also active by virtue of both their proximity to the current event and how metrically similar they are to the note

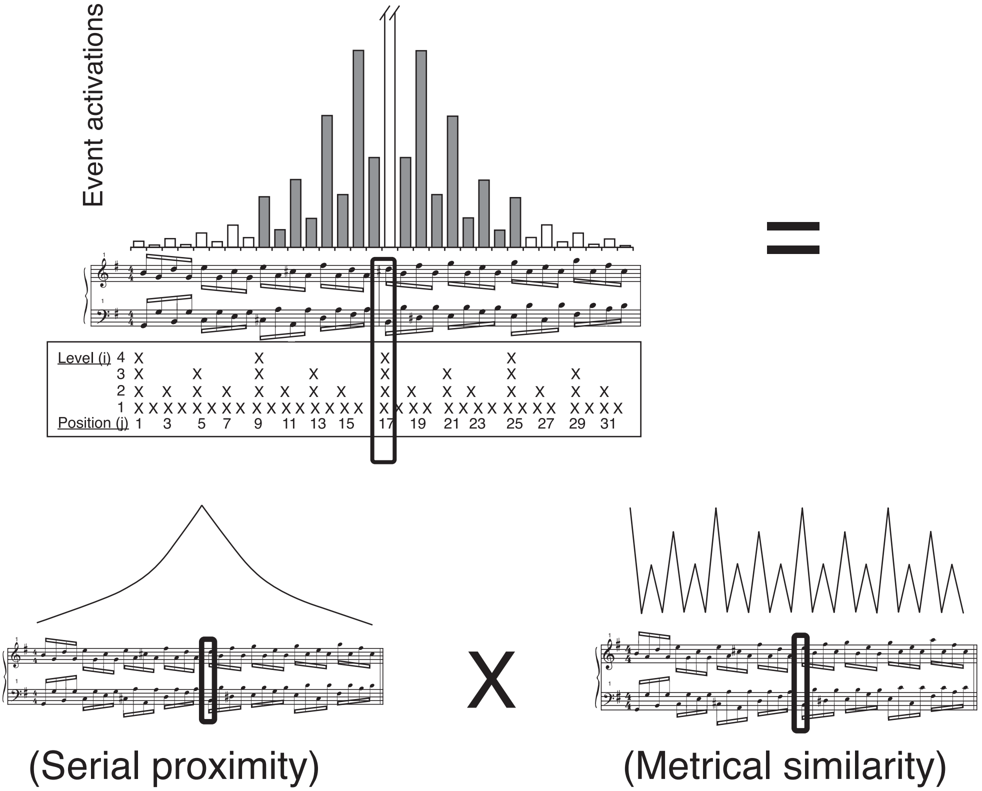


Figure 11.3 Predictions of the range model (top) along with predictions of the serial component (lower left) and metrical component (lower right).

Source: Adapted from Palmer and Pfordresher (2003, Figures 3 and 5, pp. 689 and 691) with permission from the authors and the American Psychological Association.

in the current position. What this means for the performer is that if he or she makes an error when intending to perform the D♯ note in the right hand (at the highlighted position), the intruder is more likely to be F♯ (two events in the future) or D♮ (two events in the past) than it is to be either B or A (one event in the future or past, respectively). These predictions have been verified for piano performance from memory, from notation, at different tempos (speeds), and for different age groups (Palmer & Pfordresher, 2003; Pfordresher, Palmer & Jungers, 2007).

Although it is always preferable to play accurately than to produce an error, certain serial ordering errors suggest that planning still works well. In particular, anticipations and/or errors from long distances are associated with higher skill levels and lower overall error rates than perseverations and/or errors from short distances (Drake & Palmer, 2000; Palmer & Pfordresher, 2003). Having discussed the processes involved in assembling a performance plan from memory, we now consider how performers implement this plan through motor activity. How do performers control the actions that are used to bring a performance to life?

Timing in performance

At the start of this chapter, we discussed how Glenn Gould inspired a re-thinking of Bach through his unique interpretation of the *Goldberg Variations*. This came about in large part from his use of expressive timing in the performances. We consider such use of timing in performance here, along with more basic uses of timing, such as the ability to keep a stable beat. A band member who fails to keep the beat can lead a performance to disaster. Thus, we consider timing used in such a basic way all the way to the kind of expressivity found in classic recordings.

Timing movements to match notation

In chapter 6, we described how listeners perceive the temporal structure of music. Somewhat different issues come into play when performers have to reproduce musical timing, given that the performer must communicate a rhythm to the listener. Even the seemingly simple task of keeping a regular tempo can be challenging, as anybody who has been accused of speeding up while playing understands. And when a performer successfully maintains a consistent tempo, there is always some random variability from moment to moment. Why is this?

Wing and Kristofferson (1973) developed a mathematical model, often referred to as the ***two-level timing model***, to explain this kind of timekeeping. The model was based on close analyses of timing when participants had to maintain the rate of a metronome by tapping on a table. According to this model, we keep time by coordinating two processes, both of which function imperfectly and lead to timing variability. Timing begins with the use of an ***internal timekeeper***, which reflects your brain's ability to retain a duration in memory. The internal timekeeper is effectively a mental stopwatch that keeps track of how much time has passed.

Musicians (drummers in particular) develop highly accurate internal clocks. However, no internal timekeeper is perfect. When tapping on the table, your brain will slightly overestimate or underestimate the duration from tap to tap. A common strategy for reducing the resulting variability in the internal timekeeper is to use mental subdivision. A real-world example is the way in which people subdivide numerical counting with the word 'Mississippi' so that the time increments between numbers approximate one second. The syllables of the word 'Mississippi' subdivide the time interval and allow one's timing to be more accurate. The advantage of subdividing is thought to occur because the variability of small time intervals is less than the variability of long time intervals.

The second level of the two-level model involves the interface between the internal timekeeper and muscle movements. The resulting ***motor response delay*** that occurs after the brain sends a command to the muscles can also be variable. Precision of the motor response is enhanced by physical practice. A novice musician may have a highly accurate internal sense of rhythm, but will still have trouble executing timing accurately. One intriguing prediction from this model is that when a musician switches to a new instrument that requires the use of different muscles (e.g., clarinet to drums), internal clock variability should remain consistent, but variability related to skeletal muscles should increase as the movements are not yet as highly practiced.

Assuming one can keep a steady tempo, how do performers reproduce rhythmic patterns, which are based on relative time (see chapter 6)? One idea comes from a classic theory of motor control, that of a ***generalized motor program*** (Schmidt, 1975). According to this view, your brain stores time intervals as proportional relationships, and as your rate of movement speeds up or slows down, these proportions are maintained. Note that this theory is not specific to music, but is thought to include other motor acts such as speaking and coordination of movement in sports. Some support for a role of generalized motor programs was reported by Repp (1994b), who found that recordings of performances that were sped up or slowed down in ways that maintained relative timing of inter-onset intervals sounded as natural as real performances at fast or slow tempos.

Generalized motor programs are not perfectly general, however – the idea does not apply to all forms of production. A salient exception in music comes from how jazz musicians 'swing' the beat. In principle, swinging the beat causes two equivalent quarter notes to be produced in a 'long/short' alternating pattern that upholds a 2:1 serial ratio (see chapter 6 for more on the serial ratio metric). According to the generalized motor program theory, this ratio should be maintained across all possible tempos. In practice, however, it varies greatly depending on performance tempo: Performers were found to exaggerate the ratio when playing at a slow tempo, and to reduce the ratio toward isochrony (1:1) when playing at a fast tempo (Friberg & Sundström, 2002). Even the aforementioned example by Repp (1994b) needs qualification, because certain aspects of performance timing (such as grace notes) are not timed proportionally (Desain & Honing, 1994). Ultimately, like other motor behaviors, timing in music performance is not simply determined by ratio relationships, but by other physical constraints associated with muscle movements (Gentner, 1987).

Thus, there is persistent variability in performance timing even when performers attempt to play a piece exactly as notated. But there are other sources of variability that are intended, which is what makes a performance like Gould's *Goldberg Variations* so distinctive.

Expressive variations in timing

Expressive timing refers to intentional deviations from notated timing that a performer uses to convey his or her interpretation of the music. For instance, expressive departures from strict metronomic time in the opening arpeggios of Debussy's *Arabesque* No. 1 can create a feeling of resistance during the upward swing of the arpeggio, a barely perceptible pause at the very top of the arpeggio, and a sense of release as the arpeggio descends. In the music of J. S. Bach, these departures from notated time may be subtler: a barely perceptible delay between phrases. Although these changes can be subtle, they make a big difference in how engaging a performance is for the listener. Computer notation programs can be used to generate performances with timing based precisely on music notation. These tend to sound bland and uninteresting. Expressive timing helps give life to the music, and has a considerable effect on the emotional salience of a performance, such that the listener's interpretation of how emotional the performance is can correspond with the amount of increase or decrease in expressive variations (Bhatara, Tirovolas, Duan, Levy, & Levitin, 2011).

The study of expressive timing has a long history in comparison to other topics in this chapter, partly because of the ease of measuring timing on a piano (in contrast to variations of loudness, which are harder to measure precisely). Some of the earliest empirical research on expressive timing dates back to the 1930s, in pioneering work at the University of Iowa by Carl Seashore and colleagues (Seashore, 1938/1967). In their lab, performances were analyzed and plotted on 'performance scores' that displayed variations in pitch and intensity in real time. Modern researchers have incorporated simplified versions of these representations by plotting expressive patterns above music notation, so that expressive nuances associated with particular notes align with notated durations.

Figure 11.4 shows a hypothetical example that is consistent with current literature. Suppose a performer is playing a melody at a tempo of 100 beats per minute in 4/4 time, which means that each quarter note should be held for 600 ms. A performer may choose to lengthen one note ever so slightly, say to 660 ms, based on his or her expressive intentions, or may choose to shorten it. It is important to note that these deviations from regularity in timing are distinct from random variability discussed earlier, and can be separated in statistical analyses (van Vugt, Jabusch, & Altenmüller, 2012).

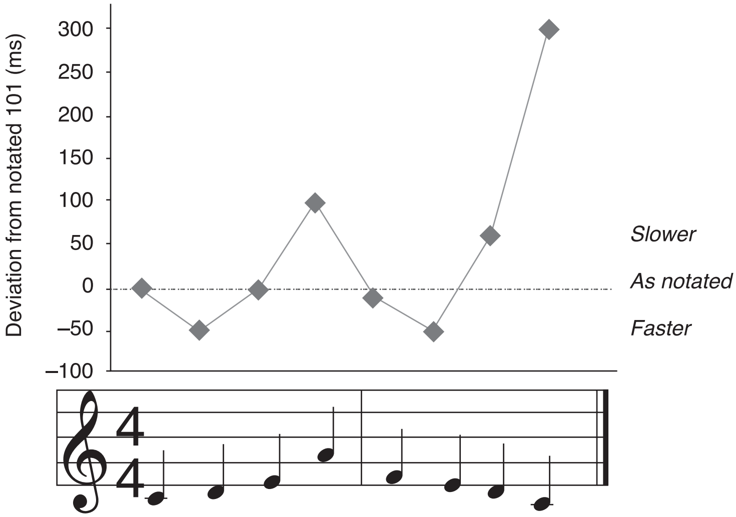


Figure 11.4 Hypothetical expressive timing of a simple melody. The music notation (below) prescribes exact inter-onset timing, but the pattern of produced inter-onset intervals (graph above) shows systematic deviations from exact regularity.

Source: Copyright © Peter Pfordresher.

Given that musical expression is often deeply personal, one might assume that the use of expressive timing is idiosyncratic, and thus not easily understood using scientific methods. In contrast to this intuition, Seashore and colleagues showed that expert performers can be very consistent in their expressive timing, a result that has often been replicated (for reviews, see Gabrielsson, 1999; Palmer, 1997). But how consistent is one performer to another? Some common trends can be found when comparing performers. One is ***phrase final lengthening***: People tend to slow down toward the end of a musical 'group.' In addition, people tend to speed up at the beginning of a phrase. The idea is similar to language. People tend to pause and slow down when they reach the end of a clause, such as in:

We have nothing to fear (pause), but fear itself!

As discussed in our chapter on rhythm perception, phrases also exist in music. For instance, consider the first four lines of the tune 'Mary Had a Little Lamb':

Mary had a little lamb,
Little lamb, little lamb,
Mary had a little lamb,
Its fleece was white as snow.

These four lines constitute two musical phrases. In musical terms, the first two lines act as a 'question' because the ending pitch does not give a sense of closure (see also chapters 5 and 7). The second two lines do provide closure based on the ending pitch, and thus act as a musical 'answer.' When performing these two lines, the performer would be likely to time each inter-onset interval in a manner that matches the phrases, with lengthening at the end of the fourth line.

The timing pattern shown in Figure 11.4 illustrates such a prototypical pattern. The lines above the notation indicate how much each note deviates from what its length would be if the person were performing exactly as the notation dictates. A 'mechanical' or 'deadpan' performance would result in all performed notes having a deviation score of zero on the Y-axis, as denoted by the dashed horizontal line on the graph. Note that this is not the same thing as plotting note durations, which could introduce variations in inter-onset intervals that are determined by the notation. In this hypothetical performance, the performer begins each phrase metronomically (a deviation of 0), speeds up to around the middle of the phrase, and then slows down near the end. Note also that the performer slows down more after the second 'answer' phrase than after the first 'question' phrase.

Some evidence for the salience of 'typical' expressive timing comes from a study in which pianists were asked to imitate patterns of expressivity after listening to a recorded performance (Clarke & Baker-Short, 1987). Some recordings had typical patterns of expressive timing, whereas others had more unpredictable ('perverse') patterns. Pianists were more successful at imitating standard patterns, in part because these recordings matched what they were naturally inclined to do themselves while performing spontaneously (Repp, 2000).

Other aspects of timing may be used to communicate the performer's interpretation. Consider, for instance, timing among notes that are notated to be played simultaneously during a piano performance. Performers typically don't perform all these notes exactly at the same time, and these slight deviations may be linked to expressive intentions. Specifically, performers often produce the melody notes slightly before the others, referred to as a ***melody lead***. Caroline Palmer (1989) has reported evidence suggesting that melody leading is used to make one of several simultaneous notes 'stand out' perceptually, typically to convey which pitch is part of the melody. In addition, evidence from perception studies shows that asynchronous events are less likely to sound as though they belong to the same Gestalt group than synchronous events (Bregman, 1990; see chapter 5 in this volume for further discussion of these principles). Timing may not be the only contributor to melody lead. The relative loudness of notes, regulated by the velocity of keystrokes, also predicts melody lead (Goebl, 2001).

Additional expressive information may be communicated by performers visually through movement (as discussed further in chapter 12). ***Motion capture*** technology, which is used in computer animation to create films such as *Avatar* and *Life of Pi*, can measure movements in real time along three dimensions. Figure 11.5 (top row) shows an example of video-based motion capture, in which movement analyses are based on extraction of image boundaries and the movement of those boundaries across time (Castellano, Mortillaro, Camurri, Volpe, & Scherer, 2008). Other forms of motion capture involve attaching small markers to the performer's joints; analyses then focus on the movement of these markers (e.g., Loehr & Palmer, 2007). An example of this technique is shown on the bottom row of Figure 11.5.

Jane Davidson (1993) used point-light displays (see bottom row of Figure 11.5) to record performers' movements while playing the piano, focusing on movements of the trunk, arms, and head. She found that performances associated with more movement are perceived as being more expressive than those with less movement (e.g., when a performer's movements are constrained). This perceived expressivity was found even when people just saw the recorded point-light displays instead of the entire film. Thus a pattern of moving dots is all the viewer has to reconstruct the underlying motion, although it is certainly more aesthetically rewarding to see the whole performer! More recently, Davidson (2005) has argued that expressive movements are hierarchically related to a common 'center of movement,' which for a pianist may be the waist. In contrast, results from a study using video-based motion capture suggest that head movements play a primary role in the communication of emotion

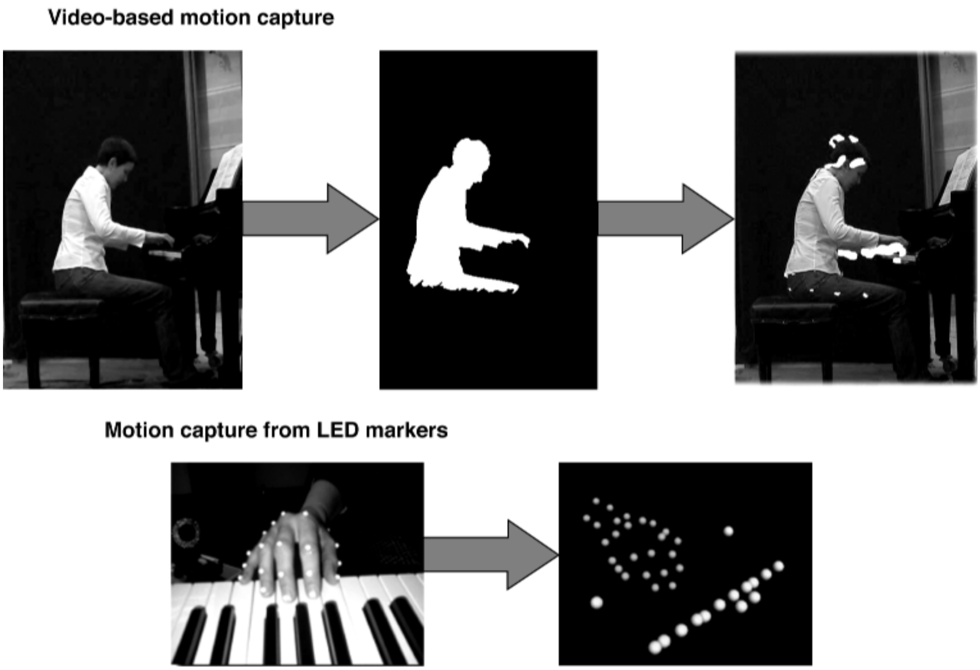


Figure 11.5 Examples of motion capture based on video (top row, Castellano et al., 2008, Figure 2) or LED markers that can be used to generate a point-light display of movement (bottom row, McGill Sequence Production Lab, www.mcgill.ca/spl).

Source: Top row reprinted with permission from the University of California Press. Bottom row © Caroline Palmer.

for piano performance (Castellano et al., 2008). In both cases, expressive movements originate in movements associated with the trunk, rather than the arms and hands.

Why do performers use expressive timing?

One reason expressive timing is used, suggested earlier, is to help communicate the structure of a piece of music to the listener. Just as pauses in speech are associated with syntax (rules for sentence structure), fluctuations in timing are linked to a kind of musical syntax. If this is the case, then it should be possible to predict expressive timing patterns across a range of performers based on the structure of a musical piece. Neil Todd (1985) proposed a mathematical model that predicts the degree of slowing a performer exhibits to the amount of 'closure' associated with a phrase. The degree of closure in his model was predicted by principles from the theory of Lerdahl and Jackendoff (1983), discussed in chapter 7. Though the predictions of this theory – which look a lot like the hypothetical example in Figure 11.4 – do not perfectly fit every situation (for detailed analyses of a number of famous performances, see Repp, 1991), it provides a starting point.

Some have observed that expressive variations in timing are reminiscent of patterns of physical motion. Expressive timing thus may act as a metaphor for movement, and may elicit the feeling of motion in the listener. For instance, Friberg and Sundberg (1999) observed that the pattern of slowing that pianists exhibit as they reach the end of the phrase resembles the pattern of slowing exhibited by runners as they slow down to a stop. There is other

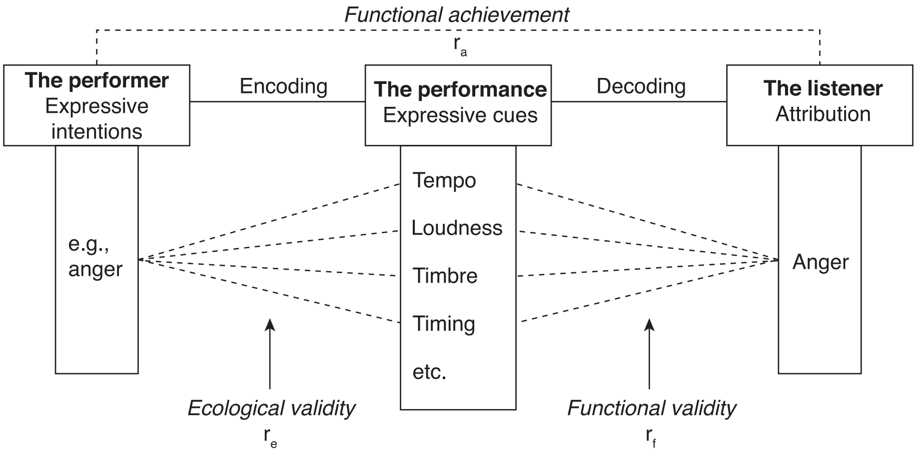


Figure 11.6 'Lens' model of emotional expression.

Source: Juslin (1997, Figure 2, p. 394), reprinted with permission from the author and University of California Press.

(somewhat controversial) evidence suggesting that music perception (particularly at high intensity levels) may engage the vestibular system, and thus give rise to perceived movement in more than a metaphorical sense (Todd & Lee, 2015). Despite the intuitive appeal of the music-as-motion analogy, this view has detractors. A specific example is a paper by Henkjan Honing (2005), who argued that expressive nuances may be better accounted for by considering the need for the performer to vary timing in ways that are perceptually salient to the listener.

The final reason expressive timing exists is perhaps the most obvious one: Expressive timing makes music more emotionally expressive. Borrowing from the Brunswikian 'lens' model, a theory from vision, Patrik Juslin (1997, 2000) proposed that the expression of emotion from performer to perceiver involves the manipulation of a set of independent expressive 'cues' by the performer. Figure 11.6 shows a schematic of Juslin's ***'lens' model of emotional expression in music***. According to this model, when the performer intends to express an emotion, he or she communicates that emotion by controlling several acoustic 'cues' that each contribute to the expressive quality of the performance. Though the performer is likely aware of the intended emotion, no awareness of the types of cues and how they are controlled is assumed. Two particularly important cues involve timing. Juslin's model refers both to variations in performance tempo and in relative timing (e.g., phrase final lengthening). The listener, when presented with these cues, can recognize an intended emotion by 'decoding' the cues, which they generally do successfully (Timmers & Ashley, 2007). Note that this model says nothing about whether an emotion is actually felt by the listener – the focus here is on the expression of emotion (see chapter 14 for more discussion of this distinction).

When giving a performance, is it better for performers to focus on their *own emotions*, induced by the music they are playing, or on communicating the emotional intent of the music? Van Zijl and colleagues asked violinists to play musical phrases in three ways: (a) *technical* performance, focusing on aspects such as correct notes, accurate rhythm, and precise articulation; (b) *expressive* performance, aiming to expressively convey the emotions in the music; and (c) *emotional* performance, focusing on the emotion they were feeling as they played (van Zijl, Toiviainen, Lartillot, & Luck, 2014). In the emotional performance condition, a sad mood was induced by giving a sad background story to the music and asking the violinist to use the emotions of the story to elicit a feeling of sadness within him- or herself. He or she would then concentrate on this sad feeling while playing. Expressivity of the performance was more exaggerated in this condition than the others, and led to a style that the authors referred to as 'extroverted,' as opposed to the more 'introverted' style of technical and expressive performances.

Tuning in performance

So far, all of the literature we have reviewed focuses on the piano. The restriction of research to the piano faces a serious limitation, however: Because pitches are represented discretely (there is no way to play a pitch in between the C and C# keys), one cannot address mechanisms used to control pitch precisely. In contrast, when one plays a violin or trombone, pitch is varied continuously, adding an additional task demand based on ***tuning***: how well a performer matches intended pitch categories. Deviations from perfect tuning may be used expressively (Morrison & Fyk, 2002). However, we focus on the basic control of tuning in singing among those who are skilled at this task as well as inaccurate or 'poor-pitch' singers. This research has been particularly influential because it broadens the focus in music performance from a small group of experts to the general population.

Singing involves three basic processes: respiration, phonation, and articulation (for an extensive review, see Sundberg, 1987). Both ***respiration*** (breathing) and ***articulation*** (moving your tongue, jaw and lips) contribute somewhat to tuning, but ***phonation*** is the primary process for controlling pitch. Inside your larynx are a set of muscles and bones that govern two flexible structures known as the ***vocal folds*** (or vocal cords). The vocal folds can be brought together (adduction) or held apart (abduction). The folds are held apart when we whisper or breathe (abduction), whereas a pitch-like quality results when the folds come together (adduction). During adduction, air from the lungs causes the vocal folds to vibrate, and the resulting frequencies of vibration generate pitch. The length, tension, and massiveness of the vocal folds are all varied when we produce different pitches. In general, high pitches are created when the folds are tenser, less massive, or (counter-intuitively) longer. Falsetto voice, for instance, is produced when the vocal folds are stretched out and thinner.

Singing is a ubiquitous musical behavior, and yet is distinctively challenging with respect to tuning. One cannot see the vocal folds without the aid of a laryngoscope, which involves inserting a tube down one's throat. By contrast, violinists can easily observe the position of their fingers to approximate what note they want to play, though most experts rely on muscle memory, proprioception (body awareness), and auditory perception. Thus it can be hard to know what physical changes are required in order to sing in tune – a predicament that has been detailed in a recent book by journalist Tim Falconer, who documents his quest to improve his own singing (Falconer, 2016).

Problems of tuning: Poor-pitch singing

We've all had the painful experience of hearing a well-intentioned friend sing a 'melody' that bears little resemblance to the intended tune. Following Welch (1979), we refer to this problem as ***poor-pitch singing***, which is a specific problem in matching the pitch of one's voice to an externally heard pitch. A surprisingly large percentage of people claim an inability to sing. In a survey of 1105 introductory psychology students, 59% claimed to be unable to imitate melodies by singing (Pfordresher & Brown, 2007). In another study, 17% of participants claimed to be 'tone deaf,' a term which they use to describe their self-assessed inability to sing, not meaning that these individuals consider themselves unable to hear pitch (Cuddy, Balkwill, Peretz, & Holden, 2005; Sloboda, Wise, & Peretz, 2005). The fact that these responses vary so much underscores the complexity of the problem. It is difficult for most people to evaluate the accuracy of their own singing. As such, neither of these figures may reflect the true prevalence of poor singing. More to the point, there are several possible deficits that may lead to poor-pitch singing.

It is, of course, possible to take the colloquial term 'tone deaf' seriously. True tone deafness would suggest that inaccurate singers have a reduced ability to hear pitch differences, resulting in a higher 'difference threshold' (the minimal pitch difference one can detect) than found in the normal population. A high enough difference threshold could interfere with singing accuracy, if for instance one were unable to distinguish different pitches of a melody. There is in fact a population who may be considered 'tone deaf,' and these individuals are generally less accurate at singing than the rest of the population (e.g., Ayotte, Peretz, & Hyde, 2002; Dalla Bella, Giguère, & Peretz, 2009; Hutchins, Zarate, Zatorre, & Peretz, 2010). The scientific term for this condition is ***congenital amusia***, meaning a lack of musical ability present throughout one's life. The primary manifestation of congenital amusia is impaired pitch perception.

Thus, there are likely to be at least some inaccurate singers who suffer from a perceptual deficit. But is this true of most bad singers? Probably not. There are some inaccurate singers who perform normally on pitch perceptual tasks (Bradshaw & McHenry, 2005). There are also people whose accuracy at reproducing pitch changes via singing outstrips their ability to discriminate pitch changes perceptually (Dalla Bella et al., 2009; Loui, Guenther, Mathys, & Schlaug, 2008).

Another hypothesis that has received some support is that poor singers have difficulty associating perceptual pitch representations with motor actions (Pfordresher & Brown, 2007), which may be the case for those with congenital amusia (Loui, Alsop, & Schlaug, 2009). Many poor singers seemed to have intact pitch discrimination ability, as well as the ability to control the pitch of their voice voluntarily, but simply are not able to link one with the other. As a result, poor singers respond differently to the sound of their own voice. Whereas artificial alterations to the shift of one's voice leads to compensatory responses in accurate singers and congenital amusics (Hutchins & Peretz, 2013), poor-pitch singers show little response to such alterations (Pfordresher & Beasley, 2014).

Finally, some poor singers may lack the ability to control the laryngeal muscles during phonation. Somewhat surprisingly, direct evidence for this kind of deficit is lacking. However, it is true that singers who are not able to translate perception into sung pitch *can* still perform similar translations in other motor systems. Hutchins and Peretz (2012) found that singers who could not match pitch vocally could match the same pitches when using a slider that could smoothly alter pitch in a continuous fashion, similar to the voice. Thus, even if singing does not involve problems within the vocal motor system, it may be based on sensorimotor connections that involve the vocal motor system.

So how prevalent is poor singing? Existing estimates vary dramatically, ranging from 13% (Dalla Bella, Giguère, & Peretz, 2007; Pfordresher & Brown, 2007) to about 60% (Hutchins & Peretz, 2012), depending on the kind of measurement authors used to divide groups. Part of the problem is that distributions of singing accuracy do not suggest two clusters, and instead suggest that singing accuracy abilities fall along a continuum (Pfordresher & Larrouy-Maestri, 2015). Furthermore, the dividing line used to separate accurate from poor-pitch singing usually depends on the context: A school choir may set a more lenient criterion for accurate singing than what one uses for Broadway auditions. What we can say at this point is that poor-pitch singing is complex both with respect to its behavioral manifestation (i.e., what kinds of pitch errors constitute poor singing) and its underlying cause. Successful pedagogical interventions for poor singers will ultimately need to be tailored to the underlying challenge that an individual confronts, and to their age (see chapter 9).

Monitoring the outcomes of performance

We now turn to the role of performer as perceiver. Upon executing a plan of action, the performer generates movements and sounds that can be perceived across multiple modalities, including sound, vision, and proprioception (perception of body position and movement). These perceived consequences serve as ***perceptual feedback*** for the actions one generates. We focus on auditory feedback, which plays a critical role in music performance and has received the most attention in the literature.

How do we use the sounds we make? The importance of having such ***auditory feedback*** present during music performance may be less critical than you think. Based on the aforementioned notion of associative chaining, one may hypothesize that hearing auditory feedback from one event is necessary to retrieve the next. If so, the absence of auditory feedback could cause a performer to 'freeze,' not knowing how to proceed without an auditory cue. This idea was in fact proposed long ago by William James (1890). Auditory feedback is not as critical for retrieval as this idea would suggest, although the presence of auditory feedback does facilitate learning. Finney and Palmer (2003) found that pianists made about 10% fewer recall errors when they memorized melodies and heard auditory feedback, compared to a condition in which memorization was carried out while playing on an electronic piano with the sound off. And as we discussed in chapter 10, neural associations between perception and action are strongly influenced by the presence of auditory feedback during learning.

However, after initial learning, the importance of auditory feedback may wane. Once a pianist has memorized a piece, removing auditory feedback while playing has only a subtle influence on performance timing, and no measurable influence on errors, when skilled pianists perform from memory (Repp, 1999). Even when playing in a duet, pianists can still synchronize with a partner when their own auditory feedback is removed (Goebl & Palmer, 2009). Of course, with a piano, one does not need to control intonation because every pitch is discretely mapped to an individual key. Is the presence of auditory feedback more important when the performer has to control intonation? Yes. Masking or removing auditory feedback has larger disruptive effects on singing (Mürbe, Pabst, Hofmann, & Sundberg, 2003) and cello playing (Chen, Woollacott, Pologe, & Moore, 2008) than on piano performance. Even so, the kind of disruptive effects found in singing and cello playing do not amount to failures of retrieval as in the associative chaining account described above. Rather, these effects are limited to the accuracy of intonation. Thus, across all these tasks, one can conclude that auditory feedback is not *necessary* for the performer to continue.

There are a couple reasons why the presence of auditory feedback is not critical after one has memorized a melody. First, as discussed in chapter 10, performers form associations between sounds and actions during learning. As a result, when a trained performer plans a sequence of actions, the brain automatically generates an imagined copy of the sequence of sounds that should result. This notion dates back again to William James, who referred to these automatic associations as ***ideomotor*** planning (James, 1890). An important implication of such action-perception associations is that performers can use ***auditory imagery*** (the conscious simulation of auditory experience; Halpern, 2003) to supplement the sounds that are absent when auditory feedback is removed (Bishop, Bailes, & Dean, 2012; Brown &

Palmer, 2012; Highben & Palmer, 2004). Research on performance errors supports this view. For instance, pianists play errors more quietly than accurate notes, suggesting that performers attempt (unsuccessfully) to halt the action that leads to the error (Repp, 1996). Brain activity during performance verifies this idea. Interestingly, the brain generates an 'error-related negativity' in the event-related potential (see chapter 4) *prior* to making the error (Maidhof, Rieger, Prinz, & Koelsch, 2009; Ruiz, Jabusch, & Altenmüller, 2009). Thus auditory feedback helps the performer monitor for errors (cf. Levelt, 1989), along with ideomotor planning.

The other reason for the weak effects of auditory feedback removal is because performers have multiple sources of perceptual information available to them. When sound is not available, performers use what they have. A convincing demonstration of this came out of the aforementioned study of piano duets by Goebl and Palmer (2009). These authors incorporated motion capture (Figure 11.4, bottom row) to track the way performers used movement as a visual cue for timing. When auditory feedback was removed, performers responded by increasing the magnitude of head and finger movements, in order to provide their partner with more salient visual cues.

Effects of altering auditory feedback

Even when the lack of auditory feedback isn't problematic, hearing the *wrong kind* of auditory feedback (e.g., by experimentally altering it) can be surprisingly detrimental. The most widely studied alteration to auditory feedback is ***delayed auditory feedback (DAF)***. As the name implies, DAF simply involves adding a temporal gap between the time at which an action is executed (e.g., a key press) and the time at which the resulting sound is expected to begin. It is like playing in a large reverberant room (e.g., a cathedral; see chapter 2 for further discussion of room acoustics), except in this room you would only hear the echo in response to your actions, and not the immediate sound produced by your instrument.

Unlike feedback absence, DAF can cause a skilled performer to sound like a beginner. Performances with DAF are characterized by slow and variable timing, and many errors (for a review, see Pfordresher, 2006). Indeed, DAF was first used with speech, and its effects on the talker were analogous to stuttering (Black, 1951; Lee, 1950). DAF's strongest effects, however, may be more on timing of actions than on the accuracy with which notes are played (Pfordresher, 2003). Ultimately, the effect of DAF, at least in music, appears to reflect the presence of disrupting rhythmic relationships between perception and action (Howell, Powell, & Khan, 1983).

Another characteristic of auditory feedback that has been investigated more recently is pitch. Interestingly, simply altering every pitch you hear in a random or random-like way while playing piano does not disrupt performance (Finney, 1997). So disruption does not result from hearing an artificially induced 'error.' However, hearing altered pitches that form a melody whose pitch structure conflicts with the pattern of pitches in the performed melody can be very disruptive. For instance, hearing a melody with the same contour but having all the upwards intervals in the melody sound when the performer is producing downwards pitch changes on the keyboard increases errors significantly, though it does not similarly disrupt timing (reviewed in Pfordresher, 2006). This kind of alteration may have an even larger effect on singing, given the greater demands of pitch control (Pfordresher & Mantell, 2012).

Although the difference between the effect of removing auditory feedback (subtle) and altering auditory feedback (subtle as a sledge hammer) may seem striking, both fit into the ideomotor account summarized above. Altered feedback disrupts performance because the sounds you hear do not match the sounds that your brain anticipates. As a result, the chain of associations between perception and action is thrown off, and the intricate coordination between the two that is necessary for performance no longer functions.

Note that auditory information from another performer can disrupt performance, just like alterations to auditory feedback. In fact, playing in a round is analogous to alterations of feedback pitch summarized above. When two performers play in a round and each member hears only the other person in the duet, the ability of both members to synchronize with each other suffers, because in this case you do need to hear yourself so that you do not confuse the sounds you are producing with their sounds (Zamm, Pfordresher, & Palmer, 2015).

Music performance and the brain

By now it should be clear that music performance involves many different cognitive processes, and thus is likely to draw on activity at many locations in the brain. For instance, Zatorre, Chen, & Penhune (2007) describe a network in which areas responsible for motor planning, particularly the primary and premotor cortices, are used to generate actions. These areas form a recurring loop with areas of the brain responsible for perception, such as the auditory cortex and the superior temporal gyrus. In addition, the premotor cortex may influence auditory areas to anticipate sounds that are consistent with planned actions. But this model only scratches the surface. In fact, all critical areas of the musical brain that are highlighted in Figure 4.2 play some documented role in music performance.

Memory retrieval during performance

Yet there are some surprises. Consider performing a complex piece of music from memory. Beyond the motoric challenges of such a performance, the demands on long-term memory are considerable. As such, one might expect higher levels of brain activity during performances of a complex memorized piece than during the performance of a musical scale, which for a trained pianist involves little or no conscious memory retrieval. But this expectation would be incorrect. A neuroimaging study using PET found *deactivations* in the frontal areas of the brain while playing a section from Bach's *Italian Concerto*, meaning that these brain areas experienced less blood flow during performances than during rest (Parsons, Sergent, Hodges, & Fox, 2005). In the study, professional pianists played some two-handed major scales and performed the Bach piece from memory while lying inside a PET scanner. Although some deactivations were also found when playing scales, these were much smaller in magnitude.

Why would one find less brain activity during a complex task like playing Bach than during a simple 'rote' task like playing scales? Parsons and colleagues (2005) speculated that these deactivations 'may be a consequence of inhibition of processes potentially able to distract the musician during a sustained performance' as well as reflecting the tendency for performers to 'lose themselves' during a performance (p. 211). In other words, it might not be good to engage in conscious reflection (frontal lobe activity) during the production of a rapid and complex motor task. Instead, it's better to let those parts of your brain that control the movements themselves (motor cortex) do the work.

These speculations adhere to a similar neural result in what might seem to be a very different task: musical improvisation. Limb and Braun (2008) found similar frontal deactivations when participants improvised jazz melodies on the piano. Although improvisation and performance of Bach differ greatly in their use of long-term memory, in both cases it is important that the performer not second-guess what to do next – a tendency that is associated with 'choking' in a variety of skilled motor behaviors (Beilock, 2010). Furthermore, it is likely that in both cases, the use of memory is procedural and non-conscious. Thus, drawing on terminology from earlier in the chapter, the neural evidence points to the use of associative chaining as opposed to content-addressable memory.

The neural basis of pitch-matching during singing

Beyond memory, the most obvious challenge of music performance is the precise and rapid control of muscles. Pianists and string players learn to coordinate finger movements in new patterns, wind instrumentalists learn to control their lip movements and breathing, and singers must learn to control laryngeal muscles to vary pitch. The ability to control these movements is governed by various brain areas, with the most fine-grained control being located in the primary motor cortex (see Figure 4.2), as discussed in chapter 10.

A particularly ambitious fMRI study of singing addressed the role of somatosensory feedback – that is, the way in which the feeling of tension in laryngeal muscles influences singing (Kleber, Zeitouni, Friberg, & Zatorre, 2013). An ear, nose, and throat physician sprayed anesthetic on the vocal folds of participants via a fiber-optic laryngoscope that was inserted through the nose. The whole process was painless – albeit unusual for participants – and the effect of anesthesia was temporary. Curiously, both singers and non-singers performed more accurately with anesthesia than without! However, the behavioral data on its own did not show the complete story. The apparent improvement from anesthesia probably reflected a compensatory response to the disruption of somatosensory feedback. Analyses of fMRI activations showed disruption within the neural network on which singing is based (a detailed account can be found in Zarate, 2013).

A particular focus of this research was the role of a brain structure called the *anterior insula*, located deep within the frontal lobe, along the fissure separating that lobe from the temporal and parietal lobes (see Figure 4.1). The anterior insula appears to play an important role in awareness of your actions, and it integrates various sensory modalities with motor actions. Surprisingly, the effect of anesthesia on the anterior insula caused opposite reactions for singers and non-singers. For singers, anesthesia led to weaker neural activations linking the anterior insula with other regions, whereas the opposite effect was found for non-singers. These differing effects may reflect the degree of flexibility found in singers versus non-singers. When confronted with a sudden block of somatosensory feedback, singers may adapt, and rely on alternate feedback sources such as auditory feedback. By contrast, less flexible non-singers may try to enhance somatosensory feedback when experiencing the numbing effects of anesthesia.

More recently, we have learned how the brain exerts fine-grained control of sung pitch. Brown and colleagues addressed this question by asking participants to engage in various tasks that used muscles of the neck and larynx, some of which involved phonation (control of pitch) and others not (Brown, Ngan, & Liotti, 2008). The tasks that involved control of pitch activated a common area of the motor cortex that the authors termed the Larynx/Phonation Area. This area was located adjacent to the area responsible for controlling the lips. Somewhat surprisingly, control of phonation was not identified in Penfield's groundbreaking work on the motor cortex, summarized in chapter 4. The discovery of a larynx area by Brown and colleagues thus constitutes a critical advance in how we understand the motor cortex. We now know of an area of the brain that can be measured to determine whether individual differences in singing ability are based on the neural control of pitch.

The neural basis of poor-pitch singing

Unfortunately, we do not know yet whether the larynx area lies behind problems of poor-pitch singing. However, several studies have addressed neural differences between congenital amusics (who are typically also poor singers) and individuals with intact pitch perception. It appears that the brains of 'tone deaf' people are structured and function differently from the brains of individuals with typical musical abilities. Structurally, the brains of congenital amusics differ from controls in the distribution of white and gray matter in the pars orbitalis region of the right frontal cortex (see Figure 4.2). Congenital amusics have proportionally less white matter (Hyde, Zatorre, Griffiths, Lerch, & Peretz, 2006) and more gray matter (Hyde et al., 2007) than controls. White matter includes neurons coated with a myelin sheath, which helps neural signals propagate more rapidly. Thus, white matter usually serves a kind of 'relay' function for information in the brain, whereas most of the 'thinking' in the brain is supported by unmyelinated gray matter. Interestingly, the pars orbitalis region is near Broca's area, and has been linked to the perception of melodic organization (Levitin & Menon, 2003; see chapter 4 in this volume).

A particularly influential paper on the brains of congenital amusics may have implications for poor singing in general. Earlier, we discussed a paper by Loui and colleagues (2008) that identified a group of congenital amusics who were able to reproduce the direction of a pitch change by singing, but were not able to report its direction verbally. In a subsequent paper, Loui and colleagues (2009) used diffusion tensor imaging (see chapter 4) to study neural connections in this group. They found that a particular structure in the right hemisphere was either greatly reduced or was missing altogether in the amusic group! The structure is known as the arcuate fasciculus, and it relays information between Broca's and Wernicke's areas of the brain (Wernicke's area is right around the planum temporale, shown in Figure 4.2).

These areas are classically known as being important for speech production and speech comprehension respectively, though we now know that they also serve more general purposes. With respect to amusia, and possibly poor singing in general, this connection may be critical for translating pitch information into actions (i.e., singing), or relating action plans for singing to one's conscious awareness of pitch. Other evidence suggests that the arcuate fasciculus is enhanced (i.e., of larger volume) and thus may be more important for singers than for nonmusicians and instrumentalists (Halwani, Loui, Rueber, & Schlaug, 2011).

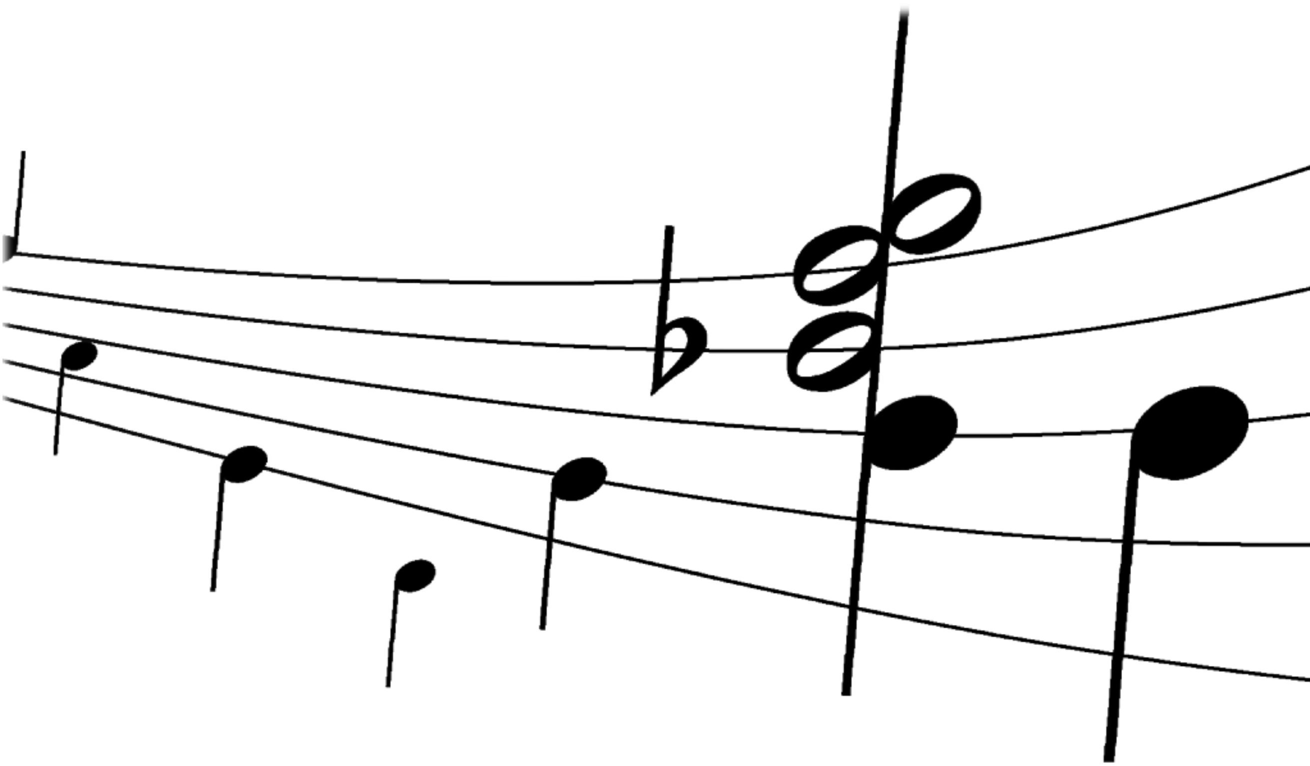
The studies selected in this review illustrate how music performance involves an intricate network of brain areas. Neuroscientific research on music has grown at a rapid pace, even since the first edition of this book, and what we offer here is just a sampling of studies that have been particularly influential and offer what we think are surprising findings. For those interested in more, we recommend the aforementioned review by Zatorre and colleagues (2007), and a more recent review by Brown, Zatorre, and Penhune (2015).

Coda

The psychology of music performance is a complex and fascinating topic. The performer must coordinate many cognitive, perceptual, and motor functions, including planning, execution, the use of feedback, and online adjustments of planning. Furthermore, performers expressively deviate from what might be considered a 'typical' performance in order to 'breathe life' into music. Performance is thus a compelling area from both a cognitive and an aesthetic perspective. Musical performance can also be studied from many other viewpoints, and the chapters that follow will examine performance in social context, as well as exploring its philosophical, emotional, and cultural significance.

Part IV

The meaning and significance of music



12 The social psychology of music

The following anecdote was recounted by Henry Fogel, retired president of the Chicago Symphony Orchestra (CSO):

> In 1990, when we were preparing at the Chicago Symphony Orchestra a set of CDs of live performances from the CSO's archives to celebrate the Orchestra's centennial, one performance that we all felt must be included was a Stokowski-led Shostakovich Tenth Symphony. As we listened to the radio tape, there was one lone person applauding after the Scherzo (which was blazingly played). Naturally, for a recording, we edited it out. But when the set was released, a few of the old-timers in the orchestra told me that we should have left it in, and explained in the accompanying booklet just who that one person was. It was Stokowski, applauding the orchestra he was conducting!
>
> (Fogel, 2007)[1]

In the opening chapters to this book, we considered sound and music at the acoustic and perceptual levels. However, they are more than simply acoustic signals and perceptual events. The sound of clapping can be heard as a mere annoyance interrupting a multi-movement musical work that has not yet concluded – a social gaffe. But on discovering that it is the applause of a world-class conductor commending his own orchestra for a spectacular performance, the sound takes on a new significance! What was initially heard as 'noise' is transformed into something of value. Similarly, as performers and listeners, we imbue the particular patterns of sound and silence that we call 'music' with deep meaning. How music takes on the significance that it does is one of its great mysteries.

The chapters in the final part of this book address this topic from a variety of perspectives to explore the social, philosophical, emotional, and cultural dimensions of music. The focus of the present chapter is on topics within the social psychology of music, and it will explore how performers and audiences influence each other, how social stereotypes may shape our appraisal of music and musicians, and how music may act as a persuasive force that shapes our behavior.

The beginnings of a social psychological perspective on music are often traced back to Paul R. Farnsworth's 1958 book, *The Social Psychology of Music*. The main aim of the book was to explore the social and cultural influences shaping musical behavior, rather than its biological bases, which Farnsworth believed had been overemphasized by the 'hereditarians' of the day. The scope of discussion on the social dimensions of music was significantly broadened several decades later in an edited book with the same title by David Hargreaves and Adrian North (1997), and more recently in another volume by the same editors *The Social and Applied Psychology of Music* (2008). These two books cover topics such as individual and social factors shaping musical taste and preference, how music influences behavior, music performance anxiety, gender and music, social factors in musical development, and applied areas such as music and health, and music in commercial contexts.

Given the diverse topics of study contained within social psychology of music, we selected a few of these topics to represent the richness of this domain. In this chapter, we will discuss how performers and audiences influence each other, the social dynamics of ensembles and orchestras, the topic of gender and music (focusing on gender stereotypes), and the influence of music on consumer behavior (as persuasion and attitude change are also central to social psychology). The variegated scope of this chapter reflects the diversity of topics inherent within this domain of psychology of music.

The power of social influence: Performers and audiences

Social psychology employs scientific methods in order to explore and understand 'how the thoughts, feelings, and behaviors of individuals are influenced by the *actual, imagined, and implied presence of others*' (Allport, 1954, p. 5, emphasis added). In this section, we explore how audiences and performers affect each other, and in the section that follows we shall examine social interaction within various performing groups.

Influence of the audience on the performers

The historical beginnings of experimental social psychology are often traced back to studies conducted in 1898 by Norman Triplett, who observed that the presence of other people tends to enhance performance, a phenomenon that later came to be referred to as 'social facilitation.' Triplett noticed that bicyclists' racing times were faster when racing against others than when racing against a clock. Even children performing simple motor tasks such as winding a fishing reel worked faster alongside a peer than when working alone. Triplett concluded that the presence of other people can enhance performance, as 'the bodily presence of another contestant participating simultaneously in the race serves to liberate latent energy not ordinarily available' (1898, p. 533).

Robert Zajonc (1965) refined the concept of social facilitation by demonstrating that the arousal evoked by the presence of others strengthens the tendency to perform 'dominant' (or highly practiced) responses, leading to performance enhancement of a well-learned or easy task, but impairing performance on a novel or complex task. One often hears of bands or ensembles that 'feed off' the energy of an audience while sounding lackluster during rehearsals. However, there are also smooth backstage rehearsals that fall flat during the public performance! Thus, the presence of an audience can lead to either ***social facilitation*** (enhanced performance in the presence of others) or ***social inhibition*** (poorer performance in the presence of others), based on the dominant response. According to Zajonc (whose name rhymes roughly with 'science'), public performances by experienced musicians performing well-worn repertoire should sparkle, whereas the same musicians venturing into an unfamiliar style of music or a newly learned complex piece may give their smoothest performances before the curtains rise!

Nickolas Cottrell (1972) further proposed that the presence of an audience is not enough to enhance the expression of the dominant response; it is their *evaluative presence* that determines the effect of performing tasks in front of others. If those present are taken to be friendly and supportive, then arousal can enhance performance, especially if the performer's task is well learned (social facilitation). However, if others present are perceived to be critical or hostile, the effects on the performance can be dire (social inhibition). This may account for one's poor performance in front of a jury, after having played the same piece flawlessly moments earlier in the practice room full of supportive peers. Further, Cottrell hypothesized that social facilitation would only take place in the presence of an audience that can evaluate one's behavior. To this end, he conducted studies demonstrating social facilitation in the presence of an audience versus alone – but not when the audience was blind-folded and unable to see if the participant performed tasks correctly (e.g., Cottrell, Wack, Sekerak, & Rittle, 1968).

Our discussion thus far has addressed the effects of the presence of others on performance in general. In the next section, we turn our attention to how audiences may affect music performance, with a focus on two topics: *performance anxiety* and *performance boost*.

'Stage fright': Music performance anxiety

When the presence of an audience brings out persistent feelings of trepidation related to performance, the condition may be referred to as 'music performance anxiety.' ***Music performance anxiety (MPA)*** refers to an enduring apprehension of musical performance that may or may not be related to one's actual level of musical ability or achievement, and may or may not have detrimental effects on performance. It is characterized by the physical sensations brought on by the activation of the sympathetic branch of the autonomic nervous system (which governs our 'flight or fight' response to danger), such as pounding heart, profuse sweating, breathlessness, trembling hands, and muscle tension, and is also manifested in cognitive, affective, and behavioral realms.

Are the pounding heart and sweaty palms always a detriment to performance? The ***Yerkes–Dodson law*** (Yerkes & Dodson, 1908) has been applied to many domains, and has been adapted to posit a relationship between level of arousal and quality of performance that can be expressed as an ***inverted-U function***, as shown in Figure 12.1. In the most basic application of this model, some degree of arousal is needed for the 'extra oomph' to give a sparkling performance, as opposed to one that may be technically proficient, but lackluster. On the other hand, whereas some arousal may enhance performance toward the peak of the inverted-U curve, too much arousal may push the performance toward the downward slope, so that quality of the performance would deteriorate.

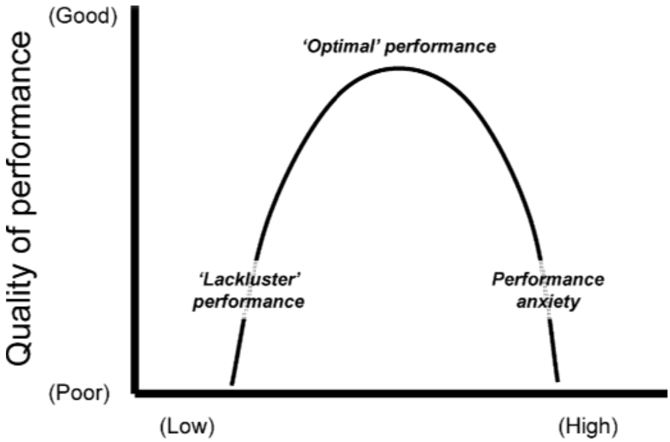


Figure 12.1 The Yerkes–Dodson law can be generally applied to performance anxiety, although 'optimal' conditions for performance will depend on many factors related to the individual performer, task, and context. Thus the height, slope, and placement of the curve will vary across individuals and performances.

In practice, of course, the relationship between arousal and performance is much more complicated and nuanced than this. What exactly constitutes 'optimal' conditions for performance depends on a complex interplay between the characteristics of the particular individual, the task, and the context (see Kenny, 2011), thus we may imagine that the placement and height and slope of the curve of the inverted-U may differ for each performer, and for different performances for the same person. For instance, returning to Zajonc's ideas on social facilitation and inhibition, well-practiced pieces of moderate difficulty level may need a higher arousal level to reach the optimal conditions for the performance to 'shine' than ones that are less practiced and either too easy or too difficult.

Some studies have shown that MPA is associated more strongly with ***trait anxiety*** (a stable disposition towards anxiety as part of one's personality make-up) than ***state anxiety*** (which is temporary arousal linked to a specific situation). Applying the Yerkes–Dodson model, an individual who possesses high trait anxiety may perform better in low-arousal contexts, while someone with low trait anxiety may benefit from greater arousal to boost his or her performance. Other personality traits linked to MPA include higher levels of neuroticism, introversion, and perfectionism (Kenny, Davis, & Oates, 2004; Sadler & Miller, 2010; Thomas & Nettelbeck, 2014). In line with Cottrell's view, a fear of negative evaluation by others seems to be at the core of music performance anxiety (Kenny, 2011; Nicholson, Cody, & Beck, 2015). Contextual variables also influence degrees of anxiety: for instance, higher levels of MPA are reported for solo than group performances, for those playing Western classical music compared to genres such as jazz or pop (Papageorgi, Creech, & Welch, 2013), and for performances in front of a live audience versus recording without an audience present, even if the performer knows the recording will be imminently released to the public via a podcast (Conklin, 2011).

How does MPA degrade performance? According to ***distraction theory*** (Eysenck, 1992), performance suffers when intrusive thoughts shift one's attention to task-irrelevant aspects (such as ruminations about what the audience may be thinking) and thus compete for limited attentional and working memory resources needed for optimal execution of the task. With diminished capacities to devote to executing the task, a formerly fluid piece can begin to wobble or an expressive performance can become mechanical, especially when playing repertoire that is fairly new or complex.

Another explanation is provided by ***explicit monitoring theory*** (Baumeister, 1984). This view predicts that performance will suffer when actions that are very well learned (or even automatized) are consciously monitored at a step-by-step level, as this interferes with the flow of automatic actions. For instance, if you begin to consciously monitor what each leg and arm must do while walking down a flight of stairs, your movements are likely to falter! Likewise, after learning a musical piece very well, conscious monitoring of the steps involved in playing (triggered by the heightened self-focus brought on by the presence of audience) can hinder the execution of well-integrated sequences of actions. While competing at a piano festival in high school, one of your authors (ST) was performing a Kabalevsky piece when – sensing the audience's attention on her – she suddenly felt acutely aware of her fingers moving over the keys and wondered, 'How do my fingers know where to go?' As she started thinking about guiding her fingers to the right keys, the performance came to a faltering stop during a piece that had been memorized months ago!

Both theories involve siphoning off resources that could be used for motor planning, efficient memory retrieval, and expressive interpretation, by occupying the performer in unproductive sorts of metacognition and heightened self-consciousness. Performers are more likely to be susceptible to 'distraction' effects when tasks are complex or not yet well learned. On the other hand, performance breakdown due to 'explicit monitoring' (as in your author's case) is more likely in highly familiarized or overlearned tasks, especially motoric tasks. These are two of many theories addressing the underlying psychological mechanisms of MPA (for a review, see Kenny, 2011, chapter 6).

Interventions for music performance anxiety

One treatment sought for performance anxiety involves administering beta-adrenaline blockers (or 'beta blockers'), as these treatments artificially reduce the performer's arousal level. Of course, such drug treatments have unfortunate side effects, one being that degree of arousal is necessary for an 'optimal' performance, as shown by the Yerkes–Dodson function in Figure 12.1. Other non-pharmacological forms of therapy focus on training the body's responses through progressive muscle relaxation, meditation, yoga, or the Alexander Technique (a method focusing on improving posture and mobility) to reduce tension and manage the physiological effects of anxiety. Among other cognitive behavioral interventions, cognitive behavioral therapy helps individuals become more aware of habitual negative thoughts, with the goal of learning to substitute them with more constructive ideas. As athlete Steve Bull once remarked: 'Nerves and butterflies are fine .... You just have to get the butterflies to fly in formation!'

More recently, virtual reality technology has been used to help performers address MPA by practicing performances in front of virtual audiences. In one pilot study, musicians performed in a room equipped with four screens and a surround-sound system that simulated the presence of a live audience – complete with people arriving and taking their seats, sounds of coughing and whispering, individuals arriving or leaving during the performance, and applause! Six one-hour sessions of virtual audience exposure significantly decreased MPA in musicians with high levels of trait anxiety (Bissonnette, Dubé, Provencher, & Moreno Sala, 2015).

Williamon, Aufegger, and Eiholzer (2014) designed an even more elaborate procedure, using a virtual environment to simulate the entire routine of a professional performance. Advanced violin students arrived in a 'green room' about half an hour before their performance time and warmed up until their stage call, waited backstage with a backstage manager, and then walked onto the stage to the sound of applause. They performed a piece in front of a simulated panel of three judges and a simulated live audience (see Figure 12.2 and note 2 for video).[2] Some violinists were also asked to perform in front of a panel of three live judges. Interestingly, no differences were found in levels of state anxiety or heart rate variability (taken with an electrocardiogram during performance) while playing in front of virtual judges versus real judges. This suggests that the simulation is potentially a useful training tool for performers, as it elicited similar responses to an actual performance. This finding resonates with Allport's aforementioned statement about how people's thoughts, behaviors, and feelings can be influenced even by the '*imagined, and implied presence of others*' (1954, p. 5). Further, the musicians in Williamon and colleagues' study reported that they believed that practicing in front of a simulated audience would not only help to manage their MPA, but also to sharpen their technical and performance presentation skills.

Some educational interventions focus on building skills that enhance the enjoyment of music-making while reducing the focus on 'right' versus 'wrong' self-evaluations. For instance, Allen (2013) showed that teaching free improvisation to children and adolescents significantly reduced their level of MPA after only six weeks of training, compared to controls who had only learned notated repertoire. Allen conjectured that this method works by

Figure 12.2 Three 'judges' (Robin Browne, William Hoyland, and Norma Jones of the Harry Partnership) on the virtual panel and a virtual audience used in the elaborate simulated audition in Williamon and colleagues' (2014) study (Figures 3A and 3B, p. 4). Will virtual environments such as these be increasingly used as a training tool for strengthening performance skills and managing performance anxiety?

emphasizing 'process over product' (as discussed in chapter 9): 'Free improvisation emphasizes the creative process, putting the decisions concerning musical content at the discretion of the performer, reducing or eliminating predetermined expectations, thus lowering levels of performance anxiety' (p. 77). This idea dovetails nicely with Cottrell's view about evaluation apprehension, as the audience may be perceived as being less able to judge an improvised performance than standard notated repertoire.

Performance boost

A few studies have focused on the potentially positive effects of performance in front of an audience (sometimes referred to as ***performance boost***). For instance, Shoda and Adachi

(2015) found that listeners preferred recordings of music performances that had been recorded in the presence of an audience, as opposed to those recorded without an audience present. Listeners gave higher ratings for technical and expressive qualities of the audience-present audio recordings than the recordings by pianists playing alone. Listeners with advanced music training also rated the audience-present audio recordings as more ‘emotionally moving’ than audience-absent recordings. Shoda and Adachi attributed the preference for audience-present recordings to Zajonc’s mere presence effect, discussed earlier: The presence of an audience seems to have boosted the pianists’ performances of a well-learned piece, adding a certain spark to the performance that listeners were able to sense from the audio recording, compared to the piano recordings produced in the solitary condition.

Musicians with lower MPA may be somewhat more likely to experience higher levels of ‘performance boost,’ characterized by positive responses to items such as ‘I feel sharper when I perform in front of an audience than when not’ or ‘Performing music gives me an extra kick when in front of an audience’ (Simoens, Puttonen, & Tervaniemi, 2015). This may be explained by drawing on the Yerkes–Dodson law (Figure 12.1). For individuals with lower MPA, the level of arousal spurred on by the presence of an audience may be sufficient to enhance the performance in an upward direction toward the peak of the inverted-U function, but would not be so great that it would push it beyond the peak to the downward slope of performance degradation.

An intriguing study suggests that a feeling of affinity with one’s musical instrument may be conducive to the enhancing effects of the presence of an audience (Simoens & Tervaniemi, 2013). In this study, 320 musicians (in voice, wind, string, keyboard, and percussion) were categorized into four groups by selecting one of the following descriptions that best characterized their relationship to their musical instrument:

(1) Person group: ‘When I perform, I feel that it’s really me as a person in front of the audience rather than my instrument/voice.’
(2) Hiding group: ‘When I perform, I feel protected/hiding behind my instrument/voice.’
(3) Obstacle group: ‘When I perform I feel that my instrument/voice is an obstacle to overcome between me and the audience.’
(4) United group: ‘When I perform, I feel so united with my instrument/voice that there is no difference between us.’

Compared to the other groups, the researchers found that those in the ‘united’ group reported lower levels of both general MPA and a more debilitating kind of MPA, and higher levels of ‘performance boost’ as assessed by a series of validated scales. However, relatively few studies have examined the individual factors and environmental conditions that may bring about the positive effects of the presence of an audience. Further research is needed to explore the potentially enhancing effects of social facilitation in musical contexts.

Influence of the performers on the audience

In a general sense, any audience is a ‘crowd,’ or a group of people who come together in some limited arena of space for a certain length of time, usually with a common purpose. With respect to crowd behavior, the etiology of extreme emotions that may be observed during live music concerts has been the subject of some studies. For instance, Lempert and

Bauer (1995) interviewed 400 people who had fainted at a rock concert and identified some possible causes for loss of consciousness at these events. Physical exhaustion and lack of food, coupled with the press of the crowd, induced the kind of oxygen starvation that ended in cases of syncope (fainting) for some people. Other concert-goers who were admitted to hospital reported intense emotions and hyperventilation that were most likely simply brought on by the natural effects of rapid breathing and screaming caused by the exhilaration of being in the presence of the performers they idolize.

It is worth noting that audience hysteria is not a modern-day phenomenon. Long before 'Beatlemania' and 'Bieber fever,' there was 'Lisztomania,' a phrase coined in 1844 by German poet Heine to describe the extreme behaviors of Liszt's admirers. Walker's (1983) influential biography of Liszt describes how admirers attempted to take clippings of his hair, and collected his coffee dregs and cigar butts to wear in vials around their necks (p. 371). Saffle (1994) further describes how women wept, threw their handkerchiefs at him, and fainted during his concerts (p. 135). In these contexts, we may see the effects not only of the performer on the crowd, but of the crowd on the people within it, as these conditions are seldom brought about when listening to music in a solitary or small group setting.

The visual impact of live performance

The fact that performers are seen as well as heard is not a trivial point, as visual information brings much to a live performance that cannot be conveyed through an audio recording (Boltz, 2013). In chapter 3, we described a study showing that long-and-graceful versus short-and-swift gestures by a marimba player affected participants' perceptions of the duration of the tones when viewing a video of the performance – even though the tones were acoustically indistinguishable in length (Schutz & Lipscomb, 2007). When participants merely listened to an audio recording of the same tones, there was no perceived difference as a result of the performer making long or short gestures when striking the keys. In another study by Vines, Krumhansl, Wanderley, and Levitin (2006), audience experience of musical tension was predominantly but not wholly determined by what was heard, whereas experience of phrasing and some structural aspects of the music were influenced by what was seen. As the authors explained, 'musical equivalents of *paralinguistic gestures* (such as head movements, eyebrow raising, and postural adjustments [that are used in everyday speech]) convey a certain amount of information that *reinforces, anticipates, or augments* the auditory signal' (p. 107).

As mentioned in the previous chapter, many studies have shown that performers' movements play an important role in communicating expressive intentions and clarifying the structure of the music being performed (Davidson, 2012). This has even been found in studies in which the visual appearance of the performer is abstracted in point-light displays (Davidson, 1993; Nusseck & Wanderley, 2009; Schutz & Kubovy, 2009), which depict movements by placing dots of lights on the performer's joints against a black background (see Figure 11.5). For instance, movements of a musician's head, upper body, hands, or elbows may trace the contour of the music and 'draw out' the shape of a melody line (Davidson, 2012).

Singers' facial expressions can convey degrees of dissonance in the music, the intensity of negative or positive emotion, and whether the melody progresses in steps or wide leaps (Thompson, Graham, & Russo, 2005). Accordingly, Thompson, Russo, and Livingstone (2010) found that listeners perceived sung intervals to be larger when accompanied by videos of vocalists singing large intervals as opposed to smaller intervals, even if the audio recording was the same. We seem to listen to music not only with our ears, but also with our eyes. Individual differences may also be at play, as listeners who started musical training earlier in life are less susceptible to being distracted by incongruent visual cues (Abel, Li, Russo, Schlaug, & Loui, 2016). Singers' expressive performances may also evoke ***facial mimicry*** in observers (Chan, Livingstone, & Russo, 2013). In Chan and colleagues' study, happy singing also elicited increased activity in the *zygomaticus* major muscle region (involved in raising the corner of the lips when smiling), and sad or neutral singing elicited increased activity in the *corrugator supercilii* muscle region (engaged when furrowing the brows, when frowning or concentrating) in those watching the singers.

Dahl and Friberg (2007) found that viewers can also identify some emotions (sadness, happiness, anger) solely from the body movements of musicians – in the absence of any sound. The instruments used in the study were marimba, saxophone, and bassoon, and despite the broad differences in the way musicians must interface with percussion versus woodwind instruments, the researchers found that the performers 'share a common body language' in conveying these emotions musically (p. 16). A study also showed that perceivers can even decode singers' emotions from movements of the face and head when viewing point-light displays of singers' faces (Quinto, Thompson, Kroos, & Palmer, 2014). Accuracy varied somewhat with the emotion expressed, the singer, and the moment at which the recording was taken during singing. But it is astonishing that listeners can pick up a singer's emotions when seeing only a few points of light, tracing the movement of a few features such as the eyebrows and lips.

In his book on musical acoustics, Hall (2001, p. 401) provides an excellent example of how a live musical performance depends on much more than sound alone. When the full orchestra is playing full force at the *fortississimo* level, the sound of the woodwind section is effectively masked by the brass and strings. As far as pure acoustics are concerned, during many rousing finales, the woodwind section could just take a nap in their seats, or pack up their cases and go home! However, the visual appearance of all performers being present on stage and 'looking busy' contributes to the full impact of a climactic passage of music for the audience.

Social interaction in performing groups: Ensembles and orchestras

The coordination of performances by members of large ensembles (in such a way that the audience hears a coherent musical event) is physically complex, considering the time it takes for the sound to reach listeners from different parts of the ensemble. The process is also socially and psychologically complex, as the performers must attend to the conductor as well as their own parts, and to the other players or singers. In smaller ensembles such as rock bands, jazz groups, and classical ensembles, there is no conductor and coordination must be accomplished in some other way. In this section, we explore the complex group dynamics of musical ensembles of various structures and sizes.

Ensembles without conductors

Small musical ensembles – such as string quartets – are typically self-governing groups made up of individuals who work closely together every day for many years, often also living and traveling together on tours. Ensembles without conductors face many challenges from within and outside the group.

For instance, the classic ***string quartet*** usually comprises two violins, viola, and cello. In practice, the coordination of an ensemble involves more complex dynamics than any simple leader/follower model would suggest, involving interdependence between parts such as second violin and cello to coordinate timing (Timmers, Endo, Bradbury, & Wing, 2014).

Figure 12.3 Listeners may perceive a sung interval to be larger when accompanied by images of a vocalist singing a wide leap (as shown by our model in the first photo) than when accompanied by images of a vocalist singing a small interval (shown in the second photo). Visual information can influence our perception of what we hear, though early music training may make listeners less susceptible to this effect.

Source: Copyright © Arianne Abela.

However, with respect to a general description of roles, the first violin part is typically given the lead and is often the focus of the audience's interest; the second violin mainly supports the melody (e.g., playing in thirds or sixths to the first violin part); the viola usually also plays an inner part, but is distinguished by timbre, which gives this part a distinct identity; and the cello usually grounds the musical piece with the bass line. Thus the second violin is usually relegated to a supporting role in ensemble performances, and is the only instrument that doubles another instrument (the lead violin) and therefore often does not have a clearly distinctive musical identity within the group.

Murnighan and Conlon's (1991) study of 20 professional string quartets in England and Scotland identified three paradoxes that often emerged within these ensembles: the *paradox of leadership versus democracy* (there is often an imbalance in power as the first violinist usually plays a lead role both musically within the group and also when interfacing with the public, though many musicians join a small ensemble in order to have an opportunity for more equality); the *paradox of the second fiddle* (the second violinist is essential to the success of the group, but is usually relegated to a supportive role in the musical repertoire and audience's perception); and the *paradox of confrontation versus compromise* (expressing disagreement can disrupt the group, but may be necessary to keep the ensemble unified and energize performances, e.g., by incorporating diverse perspectives for a richer interpretation of a musical work).

It is interesting that parallels can often be found between the musical and social roles that instrumentalists play within their ensemble, contributing to the social dynamics of the group beyond musical contexts. For instance, Murnighan and Conlon (1991) found that in general, first violinists tended to take leadership, not only musically, but in administrative roles such as public relations. First violinists in the most successful ensembles tended to lead the group while also advocating democracy, whereas first violinists of the least successful ensembles avoided taking a strong leadership role and were also viewed as less competent leaders by their group. Successful ensembles tended to have second violinists who accepted their supporting status, and were members of groups who expressed their appreciation for the second violinists openly. Finally, successful quartets avoided explicit discussion of the paradoxes or tensions described earlier, but effectively managed them implicitly within their groups. The less successful quartets openly expressed more negative attitudes toward both the group and the music.

Ensembles with different structures may encounter different group dynamics. For instance, Ford and Davidson (2003) studied members of 20 professional and semi-professional wind quintets in the United Kingdom. The technique and timbres of the instruments of the ***wind quintet*** (which typically comprises a flute, oboe, clarinet, French horn, and bassoon) are more distinct than string quartets, and intragroup dynamics may be quite different from quartets due to the odd number of members in the group. Dynamics inherent in the way the music is scored afford greater opportunities for democracy within the group, such as more equality and interchangeability among musical parts (e.g., the flute is often assigned the melody, but sometimes also plays an inner part). Unlike the more rigid string quartet, there is also more flexibility in seating arrangements for wind quintets. In addition, although the flute player was most frequently named as the leader, members of wind quintets were less likely to identify a clear leader than members of string quartets in Murnighan and Conlon's (1991) study. Instead, Ford and Davidson (2003) noted that in most of the quintets, players 'took equal responsibility for leading the ensemble' (p. 62). The musical structure and social structure of the group seem to reflect each other.

As we have seen, the dynamics of a musical ensemble are shaped by a confluence of interacting variables – including those internal to the music (e.g., the particular demands of the repertoire, the musical roles relegated to each instrument or voice), those inherent in the physical structure of the group (e.g., number of members), and those stemming from influences external to the group (e.g., public perception). In addition, there are the individual attributes of members, such as temperament, leadership skills, and proficiency level (e.g., see Kemp, 1996, on temperament).

It is interesting to observe some parallels between the musical functions of instruments inherent in the structure of the music and dynamics of the ensemble, and the social roles taken up by the musicians. Are individuals with certain traits and attributes drawn to particular instruments and roles within an ensemble? Or do the instruments that musicians play within an ensemble color their fellow musicians' perceptions, leading them to view their colleagues in ways consistent with their musical roles?

The findings of a study focusing on pop/rock musicians suggest that social perception plays an important part. Cameron, Duffy, and Glenwright (2015) found only slight differences in self-described personality traits of pop and rock musicians playing different instruments (guitar, bass, drum, vocals). However, greater differences were found for players' ratings of other band members who played different instruments. For instance, drummers rated themselves as being somewhat more 'conscientious' than other instrumentalists. However, their bandmates tended to give them much higher ratings for conscientiousness, perhaps because their perceptions were influenced by the role the drummer plays within the band, which is to provide the steady and consistent metric structure to the music. Similarly, singers tended to rate themselves somewhat higher on 'extraversion' than bassists and drummers, but the researchers speculated that their band members' even stronger perceptions of the vocalists' extraverted nature may have been augmented by singers' central and highly visible role in performances. The researchers concluded that band members 'are what they play,' but while this may be somewhat true of their personalities, the attributes associated with different instrumentalists may be more clearly manifested in the realm of social perception (p. 829).

Ensembles with conductors

Expressed in general terms, the task of a conductor and an orchestra working together is to create an orderly sequence of musical events according to certain standards of correctness, propriety, and musical aesthetics. The conductor must have the knowledge and skill to create a structured *social entity* out of the group of performers, integrated in various ways with the appreciation by the audience, in order to bring about the musical performance.

Boerner and von Streit (2007) examined the effects of a leadership style referred to as ***transformational leadership***, and identified three central components: charisma, inspirational motivation, and intellectual stimulation. A conductor adopting this style of leadership does not simply enforce his or her own artistic interpretation of a musical work on the group, 'but rather conveys it as a vision that the orchestral musicians experience as intellectual stimulation and inspirational motivation' (p. 135), interpreting familiar repertoire with novel and original concepts and a visionary quality. Another important factor leading to good performance is a 'positive group mood' among the orchestral members. Boerner and von Streit's analysis of questionnaires sent to members of 22 symphonic orchestras in Germany showed that these factors do not operate independently, but that the key is in the interplay between the two. Specifically, a transformational leadership style only has a positive effect on the artistic quality of the orchestra if there is also a high positive group mood among the musicians. In turn, a positive group mood among orchestra members only benefits the artistic quality of the orchestra if the conductor strongly embodies a transformational leadership style.

Conductors' gestures

Aside from facilitating social cohesion and inspiring the group, conductors must also coordinate the timing of the performance and provide artistic direction. Many movements of a conductor's baton are directed toward different points in space that follow a pattern consistent with the meter (see chapter 6 for a discussion of meter). The point at which the conductor reaches one of these spatial targets is conveyed by a complex interplay between radius of curvature, movement velocity, and movement acceleration. Studies in which participants synchronize with point-light displays of the conductor's baton show that no single cue on its own is enough to account for synchronization behavior, although acceleration (the rate at which movement speeds up or slows down) may be the most prominent (Luck & Sloboda, 2008, 2009; see chapter 11 in this volume on point-light displays).

Vantage point can make a difference as to how well expressive movements are perceived. Specifically, when it comes to the instrumental sections of the symphonic orchestra, Wöllner & Auhagen (2008) found that the conductor's expressive gestures may be more informative when viewed from the position of the woodwind section (in front of the conductor) and first violin section (to the conductor's left) than from where cellists and double bass players sit (to the conductor's right). This may be because the orchestral conductor performs many expressive gestures with the left hand.

Conductors' gestures may have nuanced effects on performers' muscular activity through mechanisms possibly linked to the mirror neuron system (see Rizzolatti & Craighero, 2004), as discussed in chapter 10. Social psychologists have identified the ***chameleon effect***, which refers to 'nonconscious mimicry of the postures, mannerisms, facial expressions, and other behaviors of one's interaction partners, such that one's behavior passively and unintentionally changes to match that of others' (Chartrand & Bargh, 1999, p. 893). A study using motion capture technology (see chapter 11) to measure facial movements of singers shows that singers tend to round their lips when watching a choral conductor mouth a silent /u/ vowel and raise their eyebrows when the conductor raises his or her brows (Manternach, 2012a). Singers also make more small movements of their heads or shoulders when seeing a head tilt or shoulder shrug by the conductor (Manternach, 2012b). Unintended effects have also been observed. Fuelberth (2003) found that when a conductor uses tense left hand gestures (such as a clenched fist), it often has the undesirable effect of increasing vocal tension in singers. Tension, especially in the muscles in the jaw and back of the neck and shoulders, can have detrimental effects on the quality of the vocal tone, and on the health and longevity of a singer's voice.

Conductors' gestures may also influence the audience's perception of the ensemble's performance, even when the music they produce is not actually affected. For instance, when *the same musical recording* is played with a videotape showing either a highly expressive conductor or less expressive conductor, participants tend to perceive the ensemble to have played more expressively when paired with the highly expressive conductor – even though the performance is the same. This effect has been found for both instrumental ensembles (e.g., Morrison, Price, Geiger, & Cornacchio, 2009) as well as choirs (e.g., Morrison & Selvey, 2014), as evaluated by both nonmusicians as well as musically trained audiences. Conversely, highly expressive ensemble performances (with a great range of dynamics and sensitive articulation) tend to be perceived as less expressive when paired with low conducting expressivity (Morrison, Price, Smedley, & Meals, 2014).

Figure 12.4 Laura Jackson conducts the Atlanta Symphony Orchestra. Studies have shown that a conductor's expressivity can influence listeners' judgments of how expressively the ensemble played.

Source: Photograph used with permission of Laura Jackson. Copyright © Mark Clague.

Group-level coordination

Whether performing with or without a conductor, the coordination, synchrony, and shared expressivity of performing groups requires what Keller (2001) calls ***prioritized integrative attending***. This idea refers to how a well-coordinated performance requires the careful allocation of attentional resources by ensemble performers – balancing attention to one's own part *while simultaneously* tracking the unfolding performance by the group. Both require complex tasks which involve retrieving and activating musical knowledge, executing the motor actions needed to produce the music, monitoring the outcome, and mentally representing the musical episode. Occasionally, the individual performer distributes more attention to his or her own part (e.g., when an instrumental or vocal line must be distinct from the rest), and at other times, the performer favors attending to the aggregate (e.g., when blending instrumental or vocal timbre with the group).

It is not surprising to find that conductors are particularly good at such 'prioritized integrative attending,' considering the demands of coordinating a performance such as shown in Figure 12.5. Research using ERPs suggests that the brains of conductors respond to sudden acoustic events that appear in unattended spatial locations (Nager, Kohlmetz, Altenmüller, Rodriguez-Fornells, & Münte, 2003), more so than the brains of pianists. In this study, participants listened to noise bursts from an array of speakers, with instructions to attend to sounds that came from a specific location. Sounds emanating from peripheral locations attracted involun-tary attention more readily in conductors than pianists. This experimental setup is analogous to a conductor who's attending to the first and second violins for a certain passage, yet notices an error by the second chair French horn! In another study, Wöllner and Halpern (2016) found that conductors were also better than pianists at detecting temporal

Figure 12.5 Nan Washburn is Music Director and Conductor of the Michigan Philharmonic. Whereas orchestral musicians must balance attention between their own parts and the group performance as it unfolds, conductors must allocate their attentional resources among all the performers.

Source: Photograph used with permission of Nan Washburn. Copyright © Catherine Byrd.

deviations in divided attention tasks, though they did not differ in working memory capacity. Conductors outperformed pianists in detecting small deviations in timing while listening to two streams of melodies presented simultaneously, demonstrating attentional flexibility or the ability to smoothly redirect one's focus of attention to meet the demands of a dynamic task.

Music and gender

Thus far, we have focused on how performers and audiences influence each other, and on social interactions within small and large musical ensembles. However, social psychology does not only focus on how people influence each other and interact with each other, but how they *think about* and *perceive* one another. Within the social psychology of music this includes a consideration of various social stereotypes associated with several aspects of music, particularly stereotypes concerning gender and music. In this section, we turn our attention to the topic of gender stereotypes as they pertain to the selection of musical instruments, and to the perception and evaluation of competence and prestige in performance, improvisation, and composition.

Musical instruments

The pervasive stereotyping of musical instruments and performance genres as male-appropriate or female-appropriate has been widely documented in many societies, in both historical and contemporary practice (e.g., Koskoff, 2014). In 18th-century Europe, for instance, it was common for unmarried upper- and middle-class women to play music in the parlors of their homes to showcase their standing and attract suitors. However, they were usually limited to playing such instruments as the lute, harp, piano, or violin. These soft, expressive instruments allowed the performer to embody grace and other feminine virtues, and did not involve 'distorting' the face such as when playing brass instruments or striking so-called 'indiscreet' postures such as when playing the cello. Quiet instruments were viewed as most appropriate, as they were to be played in the safe enclosure of the home. Indeed, prior to the mid-1800s, it was considered improper for women to perform in public, and most orchestras would not hire women (O'Neill, 1997a).

As women slowly joined the ranks of the professional orchestra, strong views were voiced on what instruments they could play. Gustave Kerker, a prominent music director in New York, wrote in the *Musical Standard* in 1904:

> [N]ature never intended the fair sex to become cornetists, trombonists, and players of wind instruments. In the first place they are not strong enough to play them as well as men; they lack the lip and lung power to hold notes which deficiency makes them always play out of tune .... Another point against them is that women cannot possibly play brass instruments and look pretty, and why should they spoil their good looks?
>
> (pp. 217–218)

Even today, certain gender-stereotyped ideas about musical instruments abound, and are already apparent by the time children begin their first music lessons. Many research studies have quite consistently shown that Western school-age children view flutes, violins, and clarinets as suitable instruments for girls, whereas drums, trumpets, and guitars are widely perceived as instruments that are more appropriate for boys (e.g., Abeles, 2009; O'Neill & Boulton, 1996).

Even preschool children already have stereotypical ideas about what instruments boys or girls 'should' play. For instance, Marshall and Shibazaki (2012) played excerpts of audio recordings to 3- to 4-year-old children and asked them, 'Who plays this music?', and the children indicated their response by selecting a photograph from an array of male and female faces. The researchers found that 3- and 4-year-old children already have prominent gender stereotypes for some instruments: They associated flute and violin with females, and drums with males. Unlike previous studies, the guitar, clarinet, and trumpet were not strongly associated with gender, but were mediated by the style or genre of the music. For example, when listening to a guitar being played in a classical style, the children were more likely to think a female was playing, but if the guitar was played in a rock style, they tended to think the performer was male. The finding that 3- and 4-year-olds already have gender-stereotyped ideas about different musical genres is interesting in itself.

Some studies have shown that children's attitudes do seem to change when shown gender-instrument matches that counter prevailing stereotypes. However, the observed effects are modest, and changes in attitudes are not always in the intended direction. For instance, Harrison and O'Neill (2000) presented live counter-gender-stereotyped models to children, and demonstrated a minor shift in expressed preferences among both girls and boys. However, the change was mainly seen in a decrease in preference for gender-typed instruments (i.e., girls indicated less preference for piano after observing a male pianist, and boys indicated less preference for guitar after watching a female guitarist). Another study using video presentations and counter-stereotypical drawings in Australia showed that girls were more willing to consider counter-stereotypical combinations of player and instrument than were boys (Pickering & Repacholi, 2001).

Performance

When it comes to performing artists, scholars in several fields have asserted that 'male and female artists are differentially presented, celebrated, and critiqued according to gender-loaded notions of performativity, image, and prestige' (Millar, 2008, p. 430). For example, in the realm of popular music, *Rolling Stone* magazine's list of '100 Greatest Artists of All Time'[3] includes only one woman among the top 20 artists (Aretha Franklin), and only four women in the top 50: (9) Aretha Franklin, (36) Madonna, (46) Janis Joplin, and (47) Patti Smith. In the full list, less than 10 of the 100 artists are solo female artists or feature a female lead singer in a band. It is interesting to compare *Rolling Stone*'s list (compiled by a panel of musicians, writers, and music industry executives) with *Billboard*'s list of 'most successful songs, albums, and acts of the past 50 years,'[4] which is based on actual record sales and air play since the late 1950s. Of the top 100 artists on *Billboard*'s list, 35 are solo female artists or include a female in a band.

Studies on popular music listening preferences reflect a similar bias. Millar (2008) asked college students to list their top five favorite artists or bands, the top five they believed other people in society would select, and five artists whose music they had listened to most frequently in the past three months. It was found that both male and female respondents heavily favored male artists, and that the 'pro-male' bias was reflected much more strongly on the lists submitted by the male respondents. The ratios of male-to-female artists on the 'personal favorite,' 'societal favorite,' and 'recent listening' lists were 7:1, 9:1, and 5:1 for male respondents, and 5:2, 4:1, and 2:1 for female respondents, respectively. Bias with respect to how female performing artists are perceived is evident, as perceptions about which individual artists are most valued by society do not reflect the success of female artists shown by their record sales and airplay.

When it comes to group performances, the number of female members in professional classical ensembles and orchestras is increasing and the scope of instruments represented is widening. 'Blind' auditions, which allow adjudicators to assess candidates by sound alone as they remain hidden behind a screen, have played an important role in reducing gender-biased hiring (Goldin & Rouse, 2000). However, women are still in the minority in top orchestras, and especially underrepresented in brass, string bass, and percussion sections (Phelps, 2010), echoing the findings on children's gender stereotypes of musical instruments.

Prospects are somewhat brightening in musical groups in the classical tradition, whereas gender equality has much farther to go in other genres – such as jazz. For instance, in a survey of over 600 students, McKeage (2004) found that far fewer females than males are involved in playing jazz in high school or college. Further, whereas 62% of males who played jazz in high school continued to play in college, only 26% of females who played in high school were involved in jazz in college. In particular, the female jazz musicians expressed a lack of confidence in their improvisation skills.

Adapting a scale that measures attitudes toward mathematics, Wehr-Flowers (2006) found that females in jazz bands rated themselves as significantly less confident, more anxious, and as having a lower sense of self-efficacy in jazz improvisation than males. Pointing out that most studies have *not* found significant differences in jazz improvisation abilities between males and females, Wehr-Flowers suggests that 'we must then look to alternative possibilities for the gender inequality in the jazz field' (p. 347). For instance, instruments commonly included in jazz ensembles have been stereotypically labeled as 'masculine' instruments, and females may not be socialized to feel as comfortable taking part in jazz rituals such as 'showing off one's chops' as males. Further, there is an inadequate social framework to support females, as the networks through which one obtains informal training in jazz technique and advances one's career are predominantly male (McKeage, 2004).

Composition

Another area of music in which women are still vastly underrepresented is music composition. Gender-stereotyped beliefs may play a role in accounting for the small number of females considered to be eminent composers in many genres of music (Colley, North, & Hargreaves, 2003) as demonstrated in a study employing the ***Goldberg paradigm***. This method was introduced in a classic 1968 study by Goldberg, who showed that the *same* journal articles in various fields of expertise were judged more favorably when attributed to John McKay versus Joan McKay. In Colley and colleagues' (2003) study, which applied the procedure to the musical domain, contemporary music compositions were played to 64 undergraduates who evaluated them on a set of rating scales. Although participants tended to give higher ratings on scales related to musical competence when the composers were identified as Klaus Behne and Simon Healy, compared to Helena Behne and Sarah Healy, the results were not significant. However, in another condition in which a short biography was included (that was identical for all fictitious composers), higher ratings were given on some scales for compositions attributed to female composers. The authors conjectured that if outstanding biographies are provided, readers may assume the females to be exceptionally competent to have achieved a high level of accomplishment despite the odds. Thus it is possible that the evaluations were framed by their outstanding personal accomplishments rather than a social category.

In another study employing the Goldberg paradigm, 153 participants in their late adolescence were asked to evaluate six compositions in the classical, jazz, and new age traditions (North, Colley, & Hargreaves, 2003). In this study, both the names and short biographical passages about the background information and accomplishments of (fictitious) composers were provided in all cases. When asked who would be more likely to compose in each genre, participants clearly perceived jazz composition to be a masculine activity, while indicating that females were slightly more likely to compose music in classical and new age styles.

With respect to appraisal, ratings for 'artistic merit' for the jazz compositions reflected a 'pro-female bias' by female participants and a less dramatic 'anti-female bias' by male participants than in previous research. However, the *same* jazz compositions were perceived as more 'gentle' and 'soothing' if attributed to a male composer, and more 'forceful' when attributed to a female composer. North and colleagues (2003) suggest that this may imply gender-stereotyped preconceptions about what kind of music male and female composers tend to produce. If male composers are assumed to write forceful music, a mellow piece may sound more gentle and soothing when attributed to a male composer because it runs counter to stereotypical expectations. Thus, the somewhat less pronounced bias against females by males in this study employing younger participants is hopeful, but stereotypical ideas about male and female composers were still evident in their appraisals of the music.

Other studies have shown that even highly musically trained listeners are not able to determine the sex of a composer by listening to their musical works (Sergeant & Himonides, 2016), so there do not appear to be reliable differences in the characteristics of their compositions. As with other aspects of musicality reviewed in this section, social factors – rather than raw ability – seem to account for differences between male and female composers with respect to level of perceived talent, achievement, and eminence.

Music as a social force: Consumer behavior

As our chapter draws to a close, we turn to another topic that has captured the attention of social psychologists who are interested in research applied to music in commercial contexts: the role of music in shaping the perceptions, attitudes, and behaviors of consumers in retail environments.

A number of studies show that the music in one's environment can have a considerable impact on behavior, and unlike more explicit forms of persuasion such as verbal messages, it can take place without people being consciously aware that music is guiding their behavior. This was demonstrated in the domain of consumer behavior in a study by North, Hargreaves, and McKendrick (1997). As customers milled about a supermarket, music played in the background that encouraged associations with either France (e.g., a soulful accordion piece) or Germany (e.g., brass-laden *Bierkeller* music). The researchers, posing as shoppers sitting on a bench near the cash register, recorded the customers' wine purchases. They found that shoppers' wine preferences shifted toward the country suggested by the music, even though shoppers in general were not very aware of the music. Similarly, customers in a florist's shop spent more money buying flowers when love songs and romantic music were playing in the background rather than pop music or no music (Jacob, Guéguen, Boulbry, & Sami, 2009).

Other studies have shown that music can have a profound effect on the mood and behavior of people in various commercial and industrial settings. For instance, North and Hargreaves (1998) played pop music, classical music, easy listening music, or no music in a university cafeteria. Perceptions of the ambience of the cafeteria varied with the music: When pop music was playing, the cafeteria was perceived by customers as 'fun' and 'upbeat,' when classical music was playing, it was rated as more 'sophisticated' and 'upmarket,' and when easy listening music was playing, it was perceived as 'cheap' and 'downmarket.' Further, when asked how much customers would be willing to pay for a list of 14 items sold in the cafeteria, they were willing to pay more when popular music was playing than when no music or easy listening music was playing, and prepared to spend the most money on the same items when classical music was playing. Another study taking place in a university cafeteria showed that music may even affect diners' activity rate. Diners consumed food at 3.23 bites per minute in the absence of music, 3.83 bites per minute when slow music was playing over the speaker system, and 4.4 bites per minute when fast music was playing (Roballey et al., 1985).

Possible mechanisms

Researchers have pointed to many possible mechanisms that may explain the findings reviewed in the previous segment. For instance, music may affect ***arousal***: for instance, faster or louder music may heighten physiological arousal which may speed up highly practiced behaviors, so that people may walk or talk or eat at a faster pace (as observed in the study by Roballey et al., 1985; though see Bramley, Dibben, & Rowe, 2016). However, greater speed does not always translate into higher profits: Milliman (1982) found that shoppers spent about 38% more money at a supermarket when slow music (60 beats per minute) was playing as opposed to when faster background music (108 beats per minute) was playing. Shoppers circulated through the store at a pace that was about 17% slower than when faster music was playing, and this may have allowed them to browse products that they eventually purchased (though see Knöferle, Spangenberg, Herrmann, & Landwehr, 2012).

Another possible explanation relates the findings of studies on the effects of tempo of background music to the influence of ***entrainment***, or our tendency to synchronize to an external rhythm (as discussed in chapter 6). As consumers become entrained to the pulse of the background music, it sets their tempo. In general, slower tempos may maximize profits when it is advantageous for customers to linger (e.g., to order another bottle of wine at a fancy restaurant, or to browse more products in a high-end store). However, faster tempos may bring the highest returns when quick turnaround is more advantageous (e.g., in fast food restaurants).

Another explanation draws on ***priming***, and is rooted in the power of music to prime certain thoughts or associations of a musical, cultural, or personal nature. In turn, if consumers have not yet made a firm decision concerning their options, these thoughts or associations may mediate perceptions of the store's image and influence consumer's purchasing decisions – such as whether to buy French or German wine. However, human behavior is complex, and music does not always influence consumers in predictable ways, nor serve as a direct channel to increasing sales. For instance, in one field study, researchers observed the behavior of adolescent shoppers in a candy store, in which the background music alternated between Top 40 music, music from familiar cartoon programs in France (where the study was conducted), and periods without any music. Although the study was not framed in the context of a specific theory, it was inspired by studies in which ideas 'primed' by music matching some attributes of products influenced purchasing decisions. Cartoon music should evoke associations (fun, childhood) that are congruent with candy. Indeed, customers browsed for a significantly longer time in the candy shop when cartoon music was playing. However, there was no effect on the amount of candy they purchased (Le Guellec, Guéguen, Jacob, & Pascual, 2007).

A related concept is ***musical fit***, which refers to the idea that music that seems to fit the attributes or qualities of a certain item makes those specific attributes or qualities more salient to the potential consumer. For example, Yeoh and North (2010) asked participants to view 20 items while listening to either rock or classical music. Some of the items were associated with the rebellious connotations of rock music, and others were associated with the affluent connotations of classical music, as determined by an earlier pilot study. Later, participants were asked to freely recall as many items as possible, with the same rock or classical music still playing through their headphones.

The findings showed that participants who had viewed the items accompanied by rock music recalled significantly more items associated with rock music (e.g., tattoo, long hair, electric guitar) than items associated with classical music (e.g., champagne, fountain pen, expensive watch). Further, in their free recall, participants who had viewed the items while rock music was playing tended to list 'rock'-related items first, whereas those who had viewed the items with classical music typically began their lists by first recalling the 'classical'-related items. Thus, music with the same connotations as an item (i.e., with good musical fit) can prime recall for that item, making the music-related items more likely to be remembered and more easily recalled than items not associated with the music. What this implies for advertising is that musical fit may not only influence what qualities we notice about a product or brand (e.g., that it is elegant and sophisticated), but may also strengthen memory for product recall.

We have only provided a sampling of the work in the growing area of research on music and consumer behavior. It is also important to convey the complexities of this research, as there is not a simple pathway between music and its specific effects on the thoughts, attitudes, emotions, and behavior of consumers. Human responses to music are rich and varied, and music does not simply mold behavior or manipulate people like marionettes in easily predictable ways. Interested readers are referred to Shevy and Hung (2013), North, Hargreaves, and Krause (2016), and Deaville, Rodman, and Tan (in progress) for further discussion on the role of music in advertising and in commercial contexts.

Coda

The social psychology of music encompasses a variety of topics that touch on how the 'actual, imagined, and implied presence of others' (Allport, 1954, p. 5) influences the way we engage with music. Here, the process is at least bi-directional, linking performers and audience in various ways, and also performer to performer and audience members with each other. In performance groups, a conductor actively creates a momentary social order by coordinating the musicians, who in turn influence the audience by the sounds they produce and what they communicate visually to the audience in live performances. Other ensemble performances successfully unfold without a conductor. A kind of collective forms in which features of the performance itself – such as the score, and the relationships between the instrumental parts – serve as a coordinating function. Just as social groups create music, musical performances shape the social groups that they serve to bring into being. Following the tradition of standard texts on the social psychology of music, we also explored the influence of social stereotypes concerning gender on listeners' appraisal of compositions and performances, and the power of music in shaping human behavior in commercial contexts.

Notes

1 The authors are grateful to Henry Fogel, former President of the CSO, for permission to use this quotation from *On the Record: Exploring America's Orchestras with Henry Fogel*. Retrieved from www.artsjournal.com/ontherecord/.

2 A two-minute video of the whole simulation is also posted in Williamon et al.'s online research article at http://journal.frontiersin.org/article/10.3389/fpsyg.2014.00025/full.

3 See '100 greatest artists' (2010, December 2). Retrieved from www.rollingstone.com/music/lists/100-greatest-artists-of-all-time-19691231.

4 See 'Inside *Billboard*'s Greatest of All Time: The most successful songs, albums and acts of the past 50-plus years' (2015, November 12). Retrieved from www.billboard.com/articles/events/greatest-of-all-time/6760872/billboard-greatest-of-all-time-charts-top-songs-album-acts.

13 The question of meaning in music

Material things find their places in the world of human beings as much for their significance to the people of a certain culture as for their material attributes. In one sense, a cross is just two pieces of wood. In another sense, it is a religious symbol. It is the same for sequences of sounds. In one sense, they are just sounds like any other environmental noises. In another, they are often the bearers of religious significance, complex emotions, patriotic messages, dramatic story lines, vivid images, and so on. In this and other ways, musical performances have meanings. In this chapter, we examine a number of influential theories proposed by philosophers and musicologists as to the nature of these meanings and how mere sounds can have significance for certain listeners. In chapter 14, we will address many of the questions that emerge from this theoretical discussion with empirical studies.

The sciences depend on metaphors as the foundations of theories. Among many other things, the psychology of music must explain not only what kinds of meanings music can have, but how that meaning comes about. Musicologists and music critics deploy a vast range of metaphors to try to convey what they take to be the meaning of a musical performance. Psychology of music must make use of these expositions of how meaning is to be understood in the context of musical activities of all kinds.

Understanding meaning in music

In a recent review on the role of meaning in music, Ian Cross and Elizabeth Tolbert (2016) highlight two different approaches to understanding meaning in music. One approach is linked to what we will refer to as the ***referentialist*** perspective. By this view, meaning concerns the way in which a particular symbol system (musical notes, or words in a language) bear reliable associations to things outside the symbol system. This kind of meaning system is referred to as ‘semantics’ in language, and some have suggested – controversially – that music may have its own semantics (for an extreme perspective on this topic, see Bernstein, 1976). We will summarize some thinking on how this view might apply to music; though suffice it to say that most are skeptical about how referential music may be. An important implication of the referentialist perspective is that music’s meaning is to a considerable degree determined by its context, such as social context (a commercial setting, a concert, a parade, etc., as discussed in the previous chapter). A second way of thinking about music is referred to by Cross and Tolbert as the ***aesthestic*** perspective. According to this view, music exists simply to be beautiful. Though most today would reject such a limited notion of music (much music is designed to counteract standards of ‘beauty’), a second aspect of this perspective remains strong: the idea that music’s meaning extends no further than the music itself. In other words, music may not be referential in the classic sense, but may instead be ***self-referential***.

We can begin with ***program music***, or music that is formally or informally provided with ready-made interpretations and thus is self-consciously referential. Some pieces of music have descriptive titles such as the 'Surprise Symphony' or the 'Leningrad Symphony' or 'Fingal's Cave.' There are story lines that go with the music in operas and oratorios, available moment by moment in the libretto. Popular music is commonly provided with ready-made interpretations in titles and lyrics. During the 1960s and 1970s, 'concept albums' (such as *Tommy* by The Who) unified a set of songs under a common explicit theme or story. These practices suggest that music has the power to refer to something external to itself. A movement or part of a movement might refer to a battle, a storm, a landscape, or some other idea external to the music. Music does seem to have properties that we could call semantic. Using a term drawn from philosophy, we could say music has intentionality (i.e., it is *about something*). But is this enough to account for the power music has over the thoughts and feelings of the listener? How might we explain the emotional meaning of a musical work that 'moves' us without explicit references such as these?

Understanding how music can have meaning often draws on metaphors which imply that the power of music is predominantly *emotional*. This can happen in two distinct ways. Music can *cause* or *induce* a listener to actually feel the bodily sensations that are interpreted as distinct emotions such as pride, sadness, joy. Music can also *express* emotions, for instance displaying in sound what a character in an operatic scene is feeling. Musical expression can reach the level of the transcendent, as in great religious compositions, in which feelings of sanctity and awe are the essence of the musical experience. The contrast between the emotion or mood that music induces in listeners and the emotion a listener understands some musical performance to express is complicated by the fact that the listener also often gets a certain *satisfaction* from the music whatever emotion it is conveying. Furthermore, such satisfaction occurs even when nothing that could be called an emotional state is felt or understood by the listener. 'Satisfaction' has been claimed to be the most prominent aesthetic aspect of musical experience (Butler, 2004). Some theorists, such as the great musicologist Eduard Hanslick (1891/1986), have defended the idea that it is this satisfaction, an abstract aesthetic experience, that is the only universal aspect of musical meaning.

There is a parallel with the problem of emotion in literature. Words, arranged as narratives, poems, plays, and so on can both induce and express emotions. Words, properly arranged, can also lead to that peculiar 'satisfaction' that we take in a well-made object. According to Cooke (1959):

> The task facing us [as musicologists and psychologists] is to discover exactly how music functions as a language, to establish the terms of its vocabulary, and to explain how these terms may legitimately be said to express the emotions they appear to.
>
> (p. 34)

However, most musicologists and psychologists of music would be uneasy, we believe, in taking the parallel between music and language as two major meaning systems too literally. So far we have drawn attention to some similarities, but how do they differ?

Susanne K. Langer (1942) summed up the differences between music and language as systems of meaningful signs as follows:

> The analogy between music and language breaks down if we carry it beyond the mere semantic function in general, which they are supposed to share. Logically, music has not the characteristic properties of language – separable terms with fixed connotations, and

> syntactical rules for deriving complex connotations without any loss to the constituent elements. Apart from a few onomatopoeic themes that have become conventional … music has no literal meaning.
>
> (p. 232)

Whatever music is, it is certainly very different in its means of expression from a verbal language. Music does sometimes have some referential meaning, just as some classes of words do. A musical work can stand for something outside the music itself (i.e., something 'extra-musical'), just as a proper name or a description can stand for something outside the language. Nevertheless, the way music can have referential import for a listener is not the same as the way that words refer to things.

There are many well-known examples of music imitating the sound of something that is not itself music. In Beethoven's *Pastoral Symphony*, the clarinet produces a sufficiently convincing imitation of the call of a cuckoo to leave no doubt as to which bird song is intended. In Saint-Saëns' *Carnival of the Animals*, the xylophone imitates the sound of clattering bones in 'Fossils.' In other cases, nonmusical sounds are interwoven into the music. For example, author RH vividly remembers playing the clarinet in Tchaikovsky's *1812 Overture* when the maroons, large firecrackers meant to represent cannons rather too literally by actually exploding, were set off right behind the woodwind section.

However, when the program notes offer the suggestion that a certain musical object pictures this or that scene or event, it is very easy to conjure up some alternative referent. The storm in Beethoven's *Pastoral Symphony* no doubt has the dynamic structure of a thunderstorm, along with many other natural and social phenomena. It could be rowdy villagers leaving a drunken wedding party, knocking over furniture as they go. The sound of waves breaking on a shingle beach is one possible referent for a certain passage in Debussy's *La Mer*, but not the only one. It could be the wind through a vast field of ripe corn.

The titles of these pieces bend our minds towards certain interpretations with far more latitude than our knowledge of language determines the possible referents of words and phrases. Moreover, empirical work suggests that reading verbal descriptions of music's intended meaning, does not increase, and may even diminish, the emotional experience of the listener (Margulis, 2010). Even where lyrics are supplied as an indicator of the emotional message, they may lead one astray if the music and words contradict each other to convey a sort of irony. For instance, in more recent music, the British pop group The Smiths often set hyperbolically depressing lyrics to music that would otherwise communicate happiness (such as in 'Heaven Knows I'm Miserable Now,' set in a major key and at a brisk tempo). Even in vocal music, meaning may not be completely circumscribed by the lyrics.

There are also many examples of composers providing the story in a medium other than words comprising the title or lyrics. In Prokofiev's *Peter and the Wolf*, there is a text that guides the listener in interpreting the themes played by the bassoon, oboe, flute, and other instruments. In Shostakovich's *Leningrad Symphony*, we hear the German Army marching ever closer to the city, then in a great battle of national anthems the Russian theme triumphs, and the enemy marches away. The full import of the music requires that the listener has some sense of the national significance of the melodic quotations. Tchaikovsky uses the same device of quoting fragments of the French national anthem 'La Marseillaise' and the Russian Imperial Anthem in his *1812 Overture*. Paul McCartney's 'Mull of Kintyre' is a quotation of another kind, sounding very like a Hebridean lilt, the traditional music of Scotland.

Most scholars in music cognition currently believe that apart from simple association, the emotional force and the narrative meaning of a piece of music works differently from verbal reference, or denotation. Music works by creating some sort of parallelism of structure between the referent and the form of the music, its motion, tension and resolution, and so on. In Beethoven's *Sixth Symphony*, the music is patterned in the same way as a storm that develops on a humid summer day. We hear a rumble in the distance, then the full fury of the elements, and finally the thunder dies away and the sun comes out again. Essentially and strictly in terms of the orchestral activity, we are hearing a crescendo, a passage at double forte, and then a prolonged decrescendo.

Noting that experiments made with vocal music are 'entirely unreliable' as guides in this matter, Langer points out that while 'music is known, indeed, to affect pulse rate and respiration, to facilitate or disturb concentration, to excite or relax the organism, *while the stimulus lasts* … music does not ordinarily influence behavior' (1942, p. 212, emphasis in original). Though people believe that they have the feelings because they have simply been evoked by the music they are hearing, it seems much more likely that they actually have some *idea* of what the music expresses before or during the performance. The phenomenon is as much cognitive as it is auditory and affective. Both composers and performers know how to express emotions, but they do not always actually induce them in the listener, or even intend to do so. Nor do they necessarily experience the same emotions that they express; it may hinder the performance if the flautist weeps during a sorrowful passage! Nor is music a conduit between the emotional state of the composer and that of the listener. Music is a system of symbols – that is, meaningful signs – but they are, with some exceptions, nonreferential symbols.

The units of musical meaning

Are the expressive powers of music that are manifested in aesthetic and emotional qualities explicable in terms of the basic relevant sound units that are the elementary bearers of musical significance? Meaning is not likely to reside in individual notes, since a sequence of tones has musical qualities only if the sequences are perceived as a pattern, a *gestalt*, as we demonstrated in chapter 5. Nor is it likely to be found in notes or chords taken out of the contexts of key or harmonic schemes, as we have shown in chapter 7.

One sensible suggestion was made more than half a century ago by Kate Hevner (1936; later known as Kate Hevner Mueller), emphasizing the *temporal* (or time-based) nature of music:

> It is the masses of harmony, the resolutions and progressions, to which meaning attaches, rather than to the triads themselves, and it is the twists and turns of the melody around the keynote, the meaning of the melodic line, rather than certain momentary skips and intervals, which carry the suggestiveness and the meaning of music.
>
> (p. 248)

What might be the significance of these resolutions and progressions? Reflecting on what ideas are intrinsic to music, Eduard Hanslick (1891/1986) listed 'audible changes in strength, motion, and proportion; and consequently they include our ideas of increasing and diminishing acceleration and deceleration, clever interweavings, simple progressions, and the like' (p. 10). Music, says Hanslick, can neither signify a particular object of feeling nor the feeling itself:

> It can depict not love but only such motion as can occur in connection with love or any other affect, which however is merely incidental to that affect … which moment of these

> ideas [love, anger, and fear] is it that music knows how to seize so effectively? The answer: motion. … Motion is the ingredient which music has in common with emotional states. (p. 11)

How can Hevner's and Hanslick's observations be tied in with the evident meaningfulness of many if not most musical performances? How could formal structure and motion be the sources of emotional understandings and even emotional experiences? How could these formal properties underlie the way many musical events are supposed to depict places and happenings? It is to these questions that we turn our attention in the next section. Having considered some of the ways in which musical and linguistic meaning are similar and different, we now turn to some theories of meaning as proposed by several philosophers.

Theoretical accounts of meaning in representational and nonrepresentational music

Theoreticians concerned with the topic of meaning have formulated theories of meaning in music as well as for many other kinds of representations (see Harré, 1997). We will examine three accounts of what meaning might be in the context of music.

J. Hospers (1964, pp. 44–49) distinguishes between 'pure' music and 'representational' or program music, a distinction pioneered by Hanslick (1891/1986). According to Hospers, the difference is one of degree. ***Program music***, mentioned earlier, is 'defined as music which evokes in most listeners the impression of specific objects or situations' (1964, p. 44). It is partly a matter of the composer's or the improviser's intentions, since Hospers goes on to add that representational or program music is 'designed to invoke' some specific object or event. ***Pure music***, on the other hand, does not evoke any assigned specific object for the listener. However, Hospers does not offer any theory of how representational music is able to fulfill the composer's intentions. As suggested before, perhaps it is brought about by the title or program that accompanies the music. Mendelssohn called his overture *The Hebrides*, delimiting the hearable references to the scenes encountered on a sea voyage. Debussy's lyrical titles *Reflets Dans l'Eau* (Reflections in the Water), *Clair de Lune* (Moonlight), and *Jardins Sous la Pluie* (Gardens in the Rain) all impress on the performer and listener a particular set of vivid images.

J. O. Urmson (1973) takes the distinction between representational and nonrepresentational music a step further. He points out that the use of a musical phrase to refer to something outside the music (e.g., the clarinet's imitation of the cuckoo in Beethoven's *Sixth Symphony*) is asymmetrical. Within the symphony, we can say the motif represents a cuckoo, but a springtime cuckoo call in the Vienna Woods does not represent anything in the symphony, even to Beethoven. Thus this type of representation does not work in both directions. Urmson elaborates on the idea of composer's intention in a formal definition:

'A represents B' means 'A is auditorily similar to B and X intends A to represent B.'

This definition introduces the musical context indirectly. The composer has created the passage to provide a musical footing for the representational phrase or note. Again, it must surely be the program notes or the title or something else that disambiguates the possible objects represented by the music. Urmson makes this point in arguing that 'extra-musical evidence [i.e., extrinsic to the music] is normally necessary' (1973, p. 137). Familiarity with a particular musical tradition also plays a part. Beethoven, Rossini, and Mahler all 'do' storms in a recognizable way. Knowledge about a particular composer may also play a role. A symphony by Mozart in the key of G minor, for instance, may take on greater significance for the listener who knows that Mozart had an affinity for major keys and rarely wrote in minor keys; he wrote only two symphonies in minor keys (Nos. 25 and 40), and chose G minor for both.

Hospers and Urmson take for granted that knowledge of the world of nonmusical sounds, situations, events, and so on shapes the way the music can be interpreted. On this view, the aesthetic and emotional appeal of nonrepresentational music remains a mystery. Roger Scruton, on the other hand, reverses this direction of influence. According to Scruton, music reveals to us exactly what we have put into it from our own resources. These resources come from our experience of ourselves. In Scruton's (1983) words: 'It is *our experience of ourselves*, rather than a scientific representation of the world, which both prompts and explains the metaphors we apply to music' (p. 86, emphasis added). Thus our experience of movement, rhythm, and harmony should be explicable by reference to a projection of something subjective on to an auditory object, for example the ticking of the clock as we hear it, rather than the movement of the gears and escapement of the machine itself. In this way our personal experience makes possible a particular interpretation of a heard sound pattern.

On 'movement,' a key component of musical experience, Scruton (1983) remarks 'it is ourselves as agents [i.e., the sense of ourselves as originating sequences of events] – rather than any purely geometrical idea of space – which underlies our experience of musical movement' (p. 86). He explains the perception of rhythmic patterns in a similar way in the following statements: 'The perception of rhythm involves imaginative transfer of the kind involved in metaphor' (p. 90). Thus, '[w]hen we hear a rhythm we hear sounds joining to and diverging from one another . . . in a manner familiar from our knowledge of human movement' (p. 90). Rather than the music driving the dance, the dance drives the music! Before the minuets of Haydn and Mozart and the waltzes of Strauss, there were people dancing to threesome rhythms. A great deal of current popular music is performed as part of a total experience that includes rhythmic movement as an essential core.

The roots of our experience of harmony are different. 'Harmony,' says Scruton (1983) 'belongs not to the material world of sound but to the intentional [meaningful] world of musical experience' (p. 93). Harmony appears in the 'tension, transition and resolution' to surrounding chords. The matching of patterns (or isomorphisms) between the structure of emotional experience and the structured movement of a musical object is a synthesis. It occurs because we search in the musical object for something that has its origin in ourselves.

Although it is generally thought that the capacity of music to refer to extramusical events may be fairly limited – even within program music – there is some neuroscientific evidence that supports the notion that certain general semantic associations may be triggered by musical sequences (Koelsch et al., 2004). In psychological terms, a musical sequence may prime one's memory for a particular kind of word meaning (see chapter 5 for further discussion of priming). As Koelsch and colleagues put it, 'it seems plausible that certain passages of Beethoven's symphonies prime the word *hero* rather than the word *flea*' (p. 302, emphasis in original).

They put the idea to the test in an experiment in which they measured neural responses using event-related potentials (see chapter 4). The paradigm they adopted focused on 'semantic incongruities.' In language, this involves ending a sentence with a word for which the meaning conflicts with the overall context. In a pilot study, the authors identified musical sequences that were associated with certain words. They then presented sequences to a different set of listeners in which a musical or linguistic sequence (for the music, e.g., excerpts by Strauss and Schoenberg) would be followed by a target word (e.g., 'narrow').

The authors found similar neural responses when the target word did not match its context, regardless of whether the preceding context was a melody or a sentence.

None of the accounts given above is entirely adequate when considered alone. Though too schematic to count as substantial theories, they do point the way to deeper theorizing. In chapter 14, we will survey the experimental work on how music can have meaning sufficient to invoke or express definite emotions and moods. But first we must complete our preliminary foray into this difficult topic by drawing some important conceptual distinctions needed to understand current empirical research.

Cognitivist and emotivist theories of the emotional powers of music

As we pointed out at the beginning of this chapter, music and the emotions are entangled in two very different ways. A musical performance can be the cause or the occasion for the listener, and sometimes the performer, to actually experience a certain emotion. Patriotic music might induce a bodily feeling easily interpreted as national pride. A melancholy love song might literally 'bring tears to one's eyes.' However, music can also be used to express an emotion that is understood by but not induced in the listener or the performer. A powerful aria by the Queen of the Night expresses jealous rage in Mozart's *Magic Flute*, but we can be fairly sure that no one in the audience experiences jealous rage at that point in the unfolding of the story on the stage. One does not need to be Welsh to have no difficulty in understanding what is expressed by the singing of 'Hen Wlad Fy Nhadau' ('Land of My Fathers') as the Welsh rugby team takes the field, though there may be no swelling of tear ducts or racing of the pulse except perhaps for those who are Welsh. This feature of musical expression is particularly striking in opera, where the unfolding of the story requires the listener to understand the emotion overcoming a character, without necessarily sharing it. When The Beatles brood on *chagrin d'amour*, the loss of love, in 'Yesterday,' we can sympathize without empathizing.

Music does not generally induce emotions by causing people to feel the emotions it expresses. Of course, there are individual reactions, specific occasions for effects that are sometimes no more than simple conditioning. In these cases, it would be acceptable to think of the relation between hearing the music and experiencing the emotion as a case of cause and effect. A person might have acquired a habitual emotional reaction to a certain tune by the regular association with an emotion or mood, such as the patriotic pride that is evoked by a national anthem. The idea or image of a funeral is associated with a certain tempo, the ponderous beat of a funeral march. This association enriches the listener's experience when the double bass introduces the third movement of Mahler's *First Symphony* in the manner of a funeral march, though the melody turns out to be a minor-key version of 'Frère Jacques'! But the music might have been associated on just one striking occasion with an image, a feeling or scene, so that hearing the music *brings to mind* the image, the feeling, or the scene once again. It is akin to what J. B. Davies (1978, pp. 68–70) referred to as the 'Darling, They're Playing Our Tune' phenomenon, an emotional response to music through association with a specific event.

We follow Kivy (1990) in referring to the view that music *induces* emotions in listeners as the ***emotivist*** position, and to the view that music *expresses or represents* emotions which listeners comprehend as the ***cognitivist*** position. In the psychological literature, these positions have often been presented as rivals, as if we must choose one or the other. However, it is likely that music is such that we need both positions to do justice to the relation between musical experience and emotion.

The distinction between cognitivist and emotivist positions in the psychology of musical experience is complicated by the fact that there are two contrasts in play. Cognitivists say that we understand a certain musical performance to be expressing an emotional state, by a kind of 'description,' for example a matching (or isomorphism) between the structure of the music and the structure of an emotion. Emotivists say that the musical object is causing an emotional state in the listener through physiological responses. However, the emotivist theory has been proposed in two versions, related to distinctive theories in the general psychology of the emotions. ***Reductionists*** say that an emotion is just the bodily state that is caused by the musical performance, while ***nonreductionists*** hold to a semi-cognitivist view, that only as these induced physiological states are interpreted according to certain schemata are they experienced as emotions (or moods). The same distinction appears in discussions of the *means* by which music achieves the expression of emotions.

Aside from some special cases such as the conditioned responses discussed earlier, there seems little reason to dissent from the idea that emotions are the result of *interpretations* of induced physiological states. That is, the explanation of emotional experiences in relation to music is inherently cognitive. Some research in psychology favors the cognitivist view, but in a way that finds a place for induced physiological states. In brief, it seems to be the case that specific emotions are experienced through the interpretation of physiological states according to a fairly complex cluster of cognitive schemata, including the meanings of emotions as displays of social acts. For example, a display of anger can be read by a transgressor who sparked it off as a protest at some kind of offense. According to this view, the interpretative aspects of emotions are learned as local conventions, and likely to be culturally specific.

In sum, the psychologist of music is confronted with two contrasts. Are emotions essentially physiological, or are they essentially cognitive? Is the emotionality of music its power to induce emotions in listeners, or is it its power to express emotions, though not in the way that poems and stories usually do? The truth of the matter seems to lie in some judicious blending of these contrasts.

Theories of how music and emotion are related

Having addressed some theories of meaning that encompass many contexts for meaning and introduced the cognitivist–emotivist debate, we now turn in our final section to theories that more directly address the problem of how to account for the relationship between music and emotion. We draw mainly from Susanne Langer's (1942) *Philosophy in a New Key* and musicologist Leonard Meyer's (1956) *Emotion and Meaning in Music*, while also touching on philosopher Peter Kivy's and musicologist David Huron's ideas more briefly. Most research in the area of psychology of music, even in the present time, is grounded in the seminal works of Langer and Meyer. Indeed, their work anticipated the recent shift in psychological research on emotion, from accounts dominated by the physiology of bodily reactions, towards accounts that include semantic and social aspects of emotional experience and displays.

Peter Kivy's strong cognitivism

In his spirited defense of a cognitivist position against emotivism, Peter Kivy (1990) argues that not only is it not true that music induces emotions in listeners, but that it could not do so. The way sound is heard as music precludes this possibility. His theory rests on the principle that understanding anything must be relative to perceiving it 'under a description.'

Whether a figure is seen as climbing a slope or slipping back down it will depend on the story line that accompanies the perception. If the story line is 'intrepid mountaineer,' we will see the person trudging upward. But if the story line is 'initiating an avalanche,' we will see the figure as going down to disaster!

According to Kivy (1990, p. 99), 'evidence of understanding music is the ability to describe, and enjoying music is a direct function of understanding it.' He develops this thought with the claim that the level of enjoyment of some musical object is the level of available description. A description displays that of which we are aware – that is, the musicality of the piece for this listener. For example, if the listener describes the piece in terms of the subtlety of the variations invented by the composer, the music means something different from the understanding of someone who simply notes that it is in a minor key. However, it surely cannot be the case that every activated cognitive schema must be voiceable. Kivy (pp. 127–128) notes this objection without resolving it.

Kivy's theory seems to hint at a much deeper theory of how music can have intentionality. To this end, both Meyer (1956) and Langer (1942) have important insights to offer psychologists.

Leonard B. Meyer's theory of emotional meaning of music

We now turn to musicologist Leonard B. Meyer's well-known account of the source of emotion experience in listening to music. From the outset, Meyer rejected any attempt to reduce emotional experience to physiological states and processes (Meyer, 1956, p. 13). Nor is affective musical meaning to be looked for in extramusical experience (i.e., with reference to objects and events outside of the music). Thus Meyer's account, unlike some of the other views discussed earlier, does not focus on musical meaning derived from some relation to a source external to the music itself – such as a storm, lily pads on a pond, or the pain of unrequited love – but explores musical meaning in a sense that is more intrinsic to music.

Meyer's account of meaning in music is derived from the principles of Gestalt psychology (see chapter 5), particularly the ***principle of good continuation***, which stipulates that our perceptual systems prefer interpretations of stimuli based on simple structures that do not have abrupt changes. He defines 'musical meaning' as follows: 'one musical event (be it a tone, a phrase, or a whole section) has meaning because it points to and makes us expect another musical event' (Meyer, 1956, p. 35). These expectations follow from the principle of good continuation. Meyer refers to the generation of these expectations in music as music's 'intentionality.' However, he notes that whether a musical event has intentionality for a listener depends on that listener's knowledge of musical style.

Meyer (1956, p. 37) suggests a pattern of development of the apprehension of musical meaning as the listener follows the evolution of the music. *Hypothetical* meanings arise during an act of expectation, if this (i.e., the antecedent) has happened that (i.e., the consequent) is to be expected. *Evident* meanings are attributed retrospectively to the antecedent after the consequent has occurred and the relation between antecedent and consequent has been perceived. *Determinate* meanings arise out of the relationship between hypothetical and evident meanings as the music undergoes further development. In other words, we do not know what the significance of a tone or phrase is until we have heard what follows it in the musical process.

Essentially, Meyer's account rests on a general thesis that 'emotion is evoked when a tendency to respond is inhibited [resulting in a *tension*]' (1956, p. 22). However, this broad hypothesis about the sources of emotion in general needs qualifying to differentiate it to deal with the special case of music. Given Meyer's assumption that musical experience is nonreferential (i.e., its meaning does not rely on its capacity to refer to something outside itself), the source of expectations must lie in the nature of the musical performance itself, the ordered sequence of heard tones. Wherever there are the beginnings of a gestalt pattern that are not yet completed, there must be tension. Musically, this is easily identified in various forms of incompleteness, such as the tantalizing holding back from resolution to the expected tonic in classical Western music.

To Meyer, then, tensions arise from expectations aroused by patterns that in some way are experienced consciously or unconsciously as unfulfilled or incomplete, and which may be fulfilled or thwarted. This can only happen against a background of expectation of continuity. Meyer lists a number of such continuities, the violation of which arouses emotion: melodic continuity, rhythmic continuity, and the basic and universal diatonic scale. Violations of the principle of diachronicity include chromaticism and the minor mode.

In previous chapters, we have given an account of the cognitive processes that can be used to explain the phenomena of expectation and its frustration or fulfillment that Meyer describes. The Gestalt theorists, whose ideas we discussed in chapter 5 and on whose work Meyer draws, have given some strong working hypotheses of how this could happen. In turn, Meyer's views have inspired many subsequent theories. Narmour's (1990) Implication–Realization model complements and expands on Meyer's views, and is summarized in chapter 7. Another theory that expands on Meyer's work is that of Huron (2006).

Musicologist David Huron has proposed a theory of expectation referred to as ***ITPRA*** (Huron, 2006). The acronym refers to successive stages by which we anticipate events and then respond to the manifested outcomes of observed events. The stages comprise ***imagination*** (What will come?), ***tension*** (Waiting for it to come), ***prediction*** (Did it come ... or something else?), ***reaction*** (How do I feel about what came?), and ***appraisal*** (What are the implications of this outcome?). Though his is a theory of musical expectations, Huron grounds his views in evolutionary theory, and as such, it is a theory that could respond to expectations about any temporal sequence – not just music.

ITPRA is similar to the theory of Meyer in that both rest on the assertion that musical experiences derive to a great degree from the way in which our expectations are met or thwarted. But striking differences lie in the bases for these expectations. Meyer's theory (and even more so Narmour's theory) are fundamentally ***deterministic***, in that they propose basic musical forms (or archetypes) that lead fundamentally to a certain kind of expectation. By contrast, Huron focuses on the kinds of ***statistical*** regularities in music based on the probability that a given context leads to a given outcome, as opposed to deterministic certainties. In ITPRA, expectations are formed based on the kind of experience we have, and are inherently more flexible.

As an example, Meyer proposes that the tendency for a large pitch leap (e.g., an ascending octave) to be followed by a smaller interval in the opposite direction (e.g., a descending major second) is an inherent property of music. Huron, by contrast, proposes that such forms are a byproduct of a statistical property known as *regression to the mean* (von Hippel & Huron, 2000). Small intervals are more common than large intervals, and instruments tend to stay close to an average pitch level. Thus when a melodic line makes a large movement away from this average pitch, the following interval is more likely (based only on probability) to be both smaller and to go in the direction of the average.

Susanne K. Langer: Isomorphisms of structure

In everyday life, tensions are frequently unresolved, while in music, *resolution* of tensions is at the very heart of music as an experience. In ordinary life, the source of tension and the mode of resolution may be very different in general character (a quarrel may be resolved by a brisk walk in the garden). In music, however, tension and resolution make use of the very same means, the developing pattern of an ordered sequence of tones. Meyer notes that this thesis leaves open the question as to *how* musical events 'arouse and inhibit tendencies and thereby give rise to emotions' (Meyer, 1956, p. 24). It is here that Susanne K. Langer's thesis of structural isomorphism between the forms of musical objects and those of emotional and other processes comes in to complement Meyer's theory.

Philosopher Susanne Langer's (1942) theory of how music can have meaning was one of the sources of the Meyer/Narmour Expectation/Realization proposal as the basic mechanism of musical experience. Her theory bases the listener's understanding of music on isomorphisms. In Langer's sense, ***isomorphism*** refers to patterns of similarity between the structure of the musical rhythms and melodies and the structure of other cognitive, discursive, and above all emotional 'objects' to which the music reaches out. In so far as a piece of music shares a common form with a story, a scene, a storm, a battle, or a more intimate emotional episode, so the music is capable of expressing it.

This neo-Schenkerian theory bridges the gap between the formal analyses of Schenker and Lerdahl and Jackendoff set out in chapter 7, and the phenomenology of musical experience, in explaining how music can have meaning – particularly emotional meaning. Langer's theory is further developed in her study of the general features of symbolic forms, the characteristics that endow all sorts of material things with various kinds of meanings. The structure of paintings, sculptures, landscapes, even cars can be isomorphic to emotions, story lines, and so on. Think how the shape of a Jaguar car is isomorphic on the one hand to the shape of its eponymous animal, and on the other to a sense of rapid progress of both through space (Langer, 1953).

In short: Langer's answer to how music, as a nonreferential symbol or system of symbols, can express anything is that the form of a musical event can be isomorphic with the form of something nonmusical. It might be a story, a dance step, or, in the case being discussed in this chapter, an emotional experience.

Langer draws authority from several authors who seem to have anticipated this idea. For example, she alludes to Wagner's remark that music does not express this or that person's emotion, but rather what the emotion is itself. Music, says Langer (1942), is the 'formulation and representation of emotions, moods, mental tensions and resolutions – a "logical picture" of sentient, responsive life, a source of insight not a plea for sympathy' (p. 222). Music is not a language in which emotions and life events might be described, but 'expresses primarily the composer's knowledge of human feeling' (p. 221). Langer puts this in terms of a contrast between self-expression and logical expression. The only plausible theory must be based on the idea of logical expression, that is music as significant form: '[M]usical structures logically resemble certain dynamic patterns of human experience … *what music can reflect is only the morphology of feeling*' (pp. 227, 238, original emphasis). Quoting from Carroll Pratt's (1931) *The Meaning of Music*, Langer writes:

> Music has its special, purely auditory characters, that 'intrinsically contain certain properties which, because of their close resemblance to certain characteristics in the subjective realm, are frequently confused with emotions proper.' But '… these auditory characters are not emotions at all. They merely *sound* the way moods *feel* ….'
>
> (quoted in Langer, 1942, pp. 244–245, original emphases)

To complete her account, Langer introduces the concept of the 'unconsummated symbol.' In music, 'we have an unconsummated symbol, a significant form without conventional

Figure 13.1 The mutual structural mapping of music and emotion.

significance' (1942, p. 241). It can be completed in all sorts of ways. 'The power of music,' she declares, 'lies in the fact that it can be "true" to the life of feeling in a way that language cannot; for its significant forms have that *ambivalence* of content which words cannot have' (p. 243, emphasis in original).

We can represent Langer's theory of music diagrammatically as a flow of unconsummated symbols as shown in Figure 13.1. In order to express something musical or nonmusical, the structure heard in music must match the structure of what is expressed.

In this way, music can give expression to something that is not music. To establish this claim, detailed research would be needed to reveal the relevant structures of all sorts of 'external,' nonmusical objects or events. An emotional episode can have a rising tension and some kind of resolution, whether violent or calm; a storm can be anticipated by a rumble of thunder, then the flashes of lightning and the roar of the rain, and finally the thunder dies away, and all we hear is the tinkle of rain drops falling from the leaves into the pools beneath. Love affairs, duels, hunts, and so on are all structured broadly in this way.

Summing up, we can say that according to Meyer (1956), the fulfillment and frustration of expectations induce tension against a pre-established background of 'normal' development of the music we are hearing. Following Langer (1942), music can express emotions because there are structural isomorphisms between features of music and features of emotion episodes. For example, tension/relaxation/delay in the musical form matches tension/relaxation/delay in the mental state, storyline, or whatever it might be that the music expresses. Given this general theory, our task now is to use empirical studies to identify some of the ways that the structure of music can be relevantly isomorphic to the structure of an emotion, a mood, a legend, or an everyday story. This sort of research should lead to a variety of structural theories, each appropriate to its particular domain. We turn to a survey of empirical work on this topic in the next chapter.

Coda

Theoretical accounts of the intentionality of music, in particular of the emotional power of music, take us in two contrary directions. Some pieces of music, and even certain musical phrases, become *associated* with emotions in the course of life, needing no more than the psychology of conditioning to explain it. This simple link can extend from the unique experiences of individuals to the emotionology of a whole culture. However, according to several theorists whose ideas were explored in this chapter, there is a deeper dimension: *structure*. This is where the fundamental psychological power of music lies.

Though much has been learned about the kinds of meanings that music can have for listeners, including expressing and sometimes inducing emotions, many questions remain to be answered. Music is great art. Its appeal is to that part of the psyche to which any form of great art appeals. But how does this happen? What processes occur in the listener that sustain the possibility for the induction and expression of emotion? The emotional content of musical experience ranges from the most banal of personal associations to the most refined aesthetic and even mystical evocations of the very nature of the cosmos. In the next chapter, we turn to some experimental studies on the topic of emotion in music to explore these questions.

14 The emotional power of music

> New worlds opened up, or the old one suddenly became rich and beautiful. Somebody interpreted my emotions, somebody offered a possibility to express something that was richer than anything I had come across so far. The pride, the defiance, the passion, the beauty. Somebody spoke a language I understood.
>
> (Gabrielsson, 2011, p. 200 – a woman recounting the first time she heard Beethoven's 'Appassionata')

> I had just bought Dire Straits' second album, *Communiqué*, and everything felt right to me straight away! This music expressed most of what was inside of me! … I was mortally in love with a woman …. The album spoke to me just as intensely for almost a year until we broke up …. But during all this period I played it like somebody possessed. The music … gave me a gentle, strong, self-confidence, it gave me self-esteem, an identity, gave me a new and harmonious relation to myself.
>
> (Gabrielsson, 2011, p. 201).

This chapter opens with two vibrant accounts given by participants when asked to describe their 'strongest, most intense experience of music,' as documented in Gabrielsson's (2011) book entitled *Strong Experiences with Music*. How and why does music rouse such strong emotions in us? What gives music its affective power? How is music able to move us and make us feel consoled, restored, or even transformed?

Research in the area of emotion in music has seen rich growth in the last couple of decades. In 2001, Patrik Juslin and John Sloboda published *Music and Emotion*, an edited volume that they described as 'the first scientific anthology ever devoted specifically to musical emotions' (p. 6). Less than 10 years later, the same editors noted the 'rapid development' and 'flurry of research' in the years that had ensued, in the introduction to their much-expanded *Handbook of Music and Emotion* in 2010 (p. 4). Indeed, the growth of this area is reflected in the length of the second book, which increased from almost 500 pages to nearly 1000 pages! Reviews of research in this area have also remarked on 'the accumulated mass of music and emotion studies' (Eerola & Vuoskoski, 2013, p. 308).

In the previous chapter, we considered various competing and complementary accounts of musical meaning, including meaning of an emotional nature. We now turn to some of the intriguing classic and current experimental research exploring the fascinating topic of music and emotion – addressing fundamental questions about how we perceive emotion in music, how music evokes emotions in us, and how listeners use music to regulate their emotions. The topic of cross-cultural research in this area will be discussed in our final chapter, devoted to music and culture. At the end of the current chapter, we turn to a real-world application in which the capacity of music to express and to induce emotion is particularly powerful: the role of music in film.

Emotions: Perceived and felt

Colloquially, the terms 'emotion' and 'mood' are often used interchangeably, but the standard definitions within psychology (though overlapping) are not interchangeable. In particular, ***emotions*** are thought to include feelings (a subjective experience, e.g., sorrow), and also to involve appraisals of a situation (e.g., one's cat just died), to be focused on a specific object (the cat), to be associated with physical expressions of the emotion (crying), and to be more intense but shorter in duration than 'moods.' Researchers have argued that music communicates 'emotions' in the broader psychological sense (e.g., Juslin & Västfjäll, 2008), although others have proposed that music elicits strong feelings but only generates emotions through extramusical associations (e.g., Konečni, 2008). For instance, when a song is used as a couple's wedding song, there will be a particular association with that song that may not have been there before. It is the association with that event or memory that may elicit emotions in a way that goes beyond music itself (and thus the emotional meaning is 'extramusical').

An important distinction in the research on emotion in music centers on perceived versus felt emotion. Adopting Juslin and Sloboda's (2010, p. 10) definitions, *emotion perception* refers to 'all instances where a listener perceives or recognizes emotions in music … without necessarily feeling an emotion himself or herself.' On the other hand, *emotion induction* 'refers to all instances where music evokes an emotion in a listener' and thus produces a *felt emotion*. In the first case, one might say, 'This music *sounds* sad and wistful' (***perceived emotion***), and in the other, 'This music *makes me feel* sad and wistful' (***felt emotion***). Some studies have yielded different patterns of responses for perceived versus felt emotions (e.g., Kawakami, Furukawa, Katahira, Kamiyama, & Okanoya, 2013; Zentner, Grandjean, & Scherer, 2008), suggesting that they are not the same.

With this key distinction in mind, the present chapter is organized into two main sections: (1) *perceived emotion* (highlighting studies in which instructions and tasks focused on some variation of 'What emotion did this music *express* or *convey* to you?') and (2) *felt emotion* (synthesizing studies on which instructions and tasks employed some derivative of 'how did this music make you *feel*?'). We first turn our attention to the topic of perceived emotion in music.

Perceived emotion: Emotions expressed by music

Elements of expression

One way to examine the emotions people perceive in music is to look for patterns of consistent relationships between features of music and the emotions they seem to express, as reported by listeners. This would enable us to address such questions as: What specific features of music convey or express different emotions? And do listeners generally agree on the emotions expressed by a piece of music, or are their interpretations highly variable and idiosyncratic?

A pioneering study along these lines was conducted by psychologist Kate Hevner (later known as Kate Hevner Mueller), whose ideas were mentioned in the previous chapter. Hevner (1936) focused on specific elements of music, such as major and minor key, 'firm' or 'flowing' motion in rhythm, and 'simplicity' and 'complexity' of harmony. For each element, two versions of the same piece of music were created (e.g., the same composition played in major or minor key or mode). Participants were presented with only one version of each pair, and were asked to indicate what the music expressed to them by selecting items from a list of 66 adjectives (see Figure 14.1).

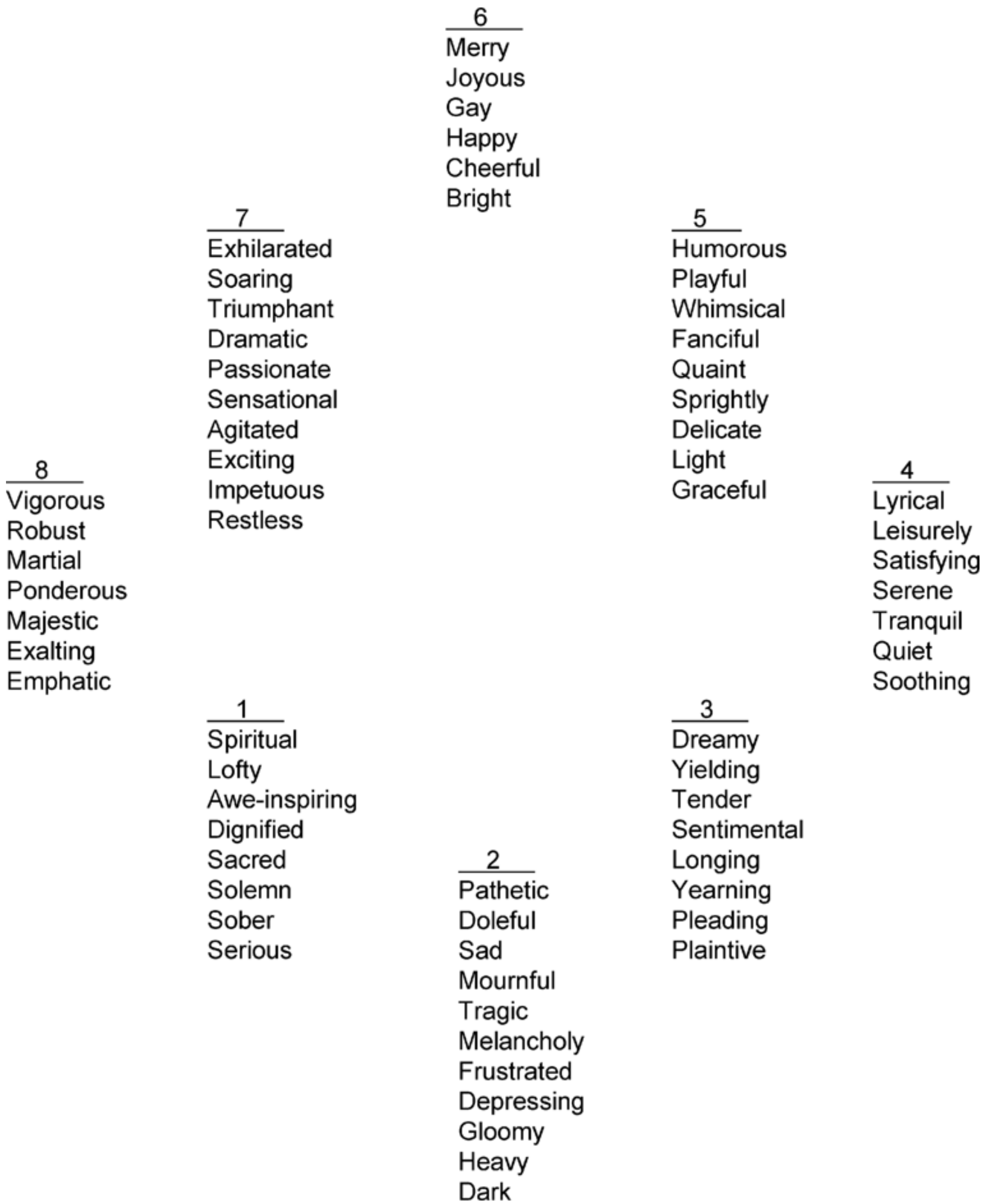


Figure 14.1 Semantic scheme used by Hevner (1936, p. 249).

Source: *American Journal of Psychology*. Copyright © 1936 by the Board of Trustees of the University of Illinois. Reproduced with permission of the University of Illinois Press.

There was general agreement among both musically trained and untrained listeners as to the affective character expressed by these musical features. The significant findings are summarized here (Hevner, 1936, p. 268):

(1) The major mode is 'happy, merry, graceful, and playful,' whereas the minor mode is 'sad, dreamy and sentimental.'
(2) Firm rhythms (with a note or chord played on every beat) are 'vigorous and dignified.' Flowing rhythms (with graceful broken chords in the accompaniment, making the motion more fluid) are 'happy, graceful, dreamy and tender.'
(3) Complex dissonant harmonies are 'exciting, agitated, vigorous, and inclined toward sadness,' whereas simple consonant harmonies are 'happy, graceful, serene, and lyrical.'

The role of tempo and mode (major or minor)

Hevner's research laid the foundations for understanding how emotion is expressed through specific features of music. Many studies have corroborated her finding that there is a high level of agreement among listeners' characterizations of some emotions expressed by a piece of music such as *happiness*, *sadness*, *anger*, and *tenderness*, whereas lower agreement has been found between listeners for other emotions such as *jealousy*, *pity*, and *whimsy* (see Eerola & Vuoskoski, 2013; Juslin, 2013b; Swaminathan & Schellenberg, 2015). Listeners draw on multiple cues to perceive emotion in music, including loudness, tempo, mode, pitch, melody, rhythm, articulation, consonance/dissonance, and timbre (Gabrielsson & Lindström, 2010), and can recognize emotions very rapidly. Surprisingly, studies have shown that both musically trained and untrained listeners can accurately perceive the intended emotion of music after only hearing *one second* of recorded music in either classical or pop/rock style (Bigand, Vieillard, Madurell, Marozeau, & Dacquet, 2005). Listeners also accurately perceive emotion after only hearing *a single chord* – even for more complex emotions such as nostalgia/longing, which are quite reliably expressed by the minor triad, minor seventh chord, and major seventh chords (Lahdelma & Eerola, 2015).

Tempo (e.g., fast or slow pace) and ***mode*** (e.g., whether the music is based on a major or minor scale) are among the most commonly studied factors in the research on perceived emotion in music. In general, music performed at a faster tempo is usually perceived as happy, and music in a slower tempo tends to be perceived as sad; pieces in a major key are usually perceived as happy, whereas pieces in a minor key tend to be perceived as sad (e.g., Gagnon & Peretz, 2003, echoing Hevner's results). Findings are especially consistent for congruent mode-and-tempo pairings, such as major-fast ('happy') and minor-slow ('sad') pieces (e.g., Hunter, Schellenberg, & Schimmack, 2010). As research on emotion in music has advanced, the scope of study has broadened to an interest in ***mixed emotions*** (i.e., music that may simultaneously express two or more emotions). For example, Hunter and colleagues played instrumental music by J. S. Bach to participants and manipulated the tempo and mode of the music independently. They found that music with conflicting cues – slow-major and fast-minor versions – led to higher ratings for mixed happy/sad emotions than music with consistent cues (fast-major, slow-minor) (Hunter et al., 2010).

These more emotionally complex responses seem to reflect listeners' engagement with music listening in their daily lives. For instance, in popular Western music, there is evidence that tastes may be gravitating toward mixed or ambiguous emotions. Schellenberg and von Scheve (2012) analyzed over 1000 Top 40 recordings from 1965 to 2009. They found that the proportion of songs in minor keys almost quadrupled between the first and last samples they analyzed: Only 15% of Top 40 songs were composed in a minor key in 1965–1969, whereas 57.5% of Top 40 songs were in a minor key in 2005–2009.

Most pertinent to our discussion, they also found that the tempo of Top 40 songs has become significantly slower since the 1960s, especially for songs in *major* keys. As slow tempo is usually associated with sadness, this is a mixed emotional cue when combined with the major mode in Western music. The researchers concluded that popular music most favored by general audiences 'became sadder sounding and more emotionally ambiguous since the 1960s' (Schellenberg & von Scheve, 2012, p. 600). Conversely, 'unambiguously happy' songs (major key, fast tempo) have become less common on Top 40 lists. The researchers speculated that the growing appreciation for emotional complexity in music may reflect broader social and cultural changes in society at large over the past decades, including the tendency to question simple categorizations in many areas (such as challenging traditional gender roles or binary concepts of gender).

Our discussion has focused on mode and tempo as a reflection of their prominence in previous research on music and emotion. Of course, mode and tempo are not the only cues that listeners rely on to perceive emotion in music. For instance, McPherson and colleagues asked jazz pianists to improvise music to convey 'positive,' 'negative,' and 'ambiguous' emotions. Listeners were then asked to listen to the audio recordings of the performances, and although about 29% of the 'positive' improvisations were mostly in a minor key and about 35% of 'negative' improvisations were mostly in a major key, listeners were able to identify the intended emotions. In these cases, other features of the music seemed to differentiate the performances: for instance, 'positive' improvisations comprised more notes played per second, faster key presses, a lighter touch or staccato, and more playing in the high range of the piano (McPherson, Lopez-Gonzalez, Rankin, & Limb, 2014) than 'negative' or 'ambiguous' improvisations.

To this point, we have focused on how music may communicate positive ('happy') or negative ('sad') emotions, as this is arguably the most salient distinction between emotions. Nevertheless, according to Russell's (1980) ***circumplex model of affect***, musical emotions may be based on two interacting dimensions: ***valence*** (the degree to which an emotion is associated with favorable outcomes, i.e., positive and negative) and ***arousal*** (the amount of energy associated with an emotion). This model allows for some more complexity and further elaboration of mixed emotions. For instance, whereas pairing a major mode with a fast tempo is predicted to produce 'happiness,' pairing a minor mode with a slow tempo leads to 'calmness' – which is a 'mixed' emotion that is positively valenced (favorable), but low in energy (Husain, Thompson, & Schellenberg, 2002).

Developmental studies on perceived emotion in music

How soon does the ability to perceive emotions expressed in music emerge during the life course? As they are not yet well acquainted with the conventions of any particular musical style, young children tend to rely on basic psychoacoustic cues such as tempo, loudness, and range of pitches to interpret emotion in music (Adachi, Trehub, & Abe, 2004), as do adults judging emotions in music from another culture (Balkwill & Thompson, 1999; see also chapter 15).

Children also put more weight on certain cues than adults, such as focusing on the lyrics to judge the emotion in a singer's performance (Morton & Trehub, 2007). For instance, when the lyrics of a song are 'I had my favorite cake for dessert,' 5- to 10-year-old children tend to label the performance as 'happy' singing, even if the performer sings in a 'sad' style that is very slow and in a minor key. However, if songs are sung on a nonsense syllable like 'da,' children are able to focus on the style of the performer's singing and the characteristics of the music to decipher the emotions expressed by the music, as most adults do. In general, children are less accurate than adults at identifying the emotions that a performer intends to express through music. However, when deciphering emotions of expressive performances by other children, they outperform adults (Adachi et al., 2004), just as children are more accurate than adults at interpreting emotional facial expressions of other children (Profyt & Whissell, 1991).

When it comes to instrumental music, 3-year-old children are sensitive to the general positive and negative connotations of music in major and minor keys (Kastner & Crowder, 1990). They can reliably point to a line drawing of a face expressing a positive emotion (happy, content) when hearing music in a major key, and point to a face expressing a negative emotion (sad, angry) when hearing music in a minor key. However, their responses are not yet nuanced enough to distinguish specific emotions. By 5 years old, most children can accurately recognize some basic emotions such as happiness or sadness expressed in music, though they often confuse fear and anger (Terwogt & van Grinsven, 1991).

Subsequent studies also show that 5-year-olds tend to perceive happiness or sadness in music based mainly on cues such as tempo (slow = sad, fast = happy), but not yet on mode (major key = happy, minor key = sad) (e.g., Dalla Bella, Peretz, Rousseau, & Gosselin, 2001). Whereas the association between tempo and emotion may be more intuitive (as discussed later in this chapter), the connotations of major and minor keys seem to require some exposure and learning (as discussed in chapter 8). As they become familiar with the pitch structures and conventions of a particular musical tradition, children gradually improve in their ability to interpret intended emotions, and perform similarly to adults in identifying emotions expressed by music by about 11 years of age (Hunter, Schellenberg, & Stalinski, 2011). Specifically, in Hunter and colleagues' study, 11-year-olds were as accurate as adults at recognizing happiness, fear, peacefulness, and sadness expressed in instrumental music.

The ability to accurately recognize positive emotions in music (such as happiness and peacefulness) appears to be quite stable across the lifespan, whereas the ability to identify negative emotions in music (such as sadness and fear) gradually declines from mid-life through old age (Lima & Castro, 2011a). Incidentally, these findings are mirrored in the research on felt emotions (i.e., emotions *induced* by the music): Older adults also report feeling more intense happiness when listening to happy music, and are less responsive to music inducing sadness and fear compared to college students (Vieillard & Gilet, 2013). These findings are often explained in terms of adaptive emotional coping mechanisms of older adults, such as the ***positivity effect*** (i.e., the enhanced processing of positive information and reduced processing of negative information) in later life, in tandem with changes in brain structure and functioning associated with advancing age.

Interestingly, long-term and intensive engagement in music may offset this age-related trend. A follow-up study did not detect this mid-life decline in perceiving negative emotions in music in musicians (Castro & Lima, 2014); musicians aged 40 to 60 years maintained their accuracy in perceiving both positive and negative emotions in music, presumably because the ability to perceive both kinds of emotions is involved in sensitive appreciation and interpretation of music.

How do we recognize emotion in music?

Given that individuals are able to consistently perceive certain intended emotions in music, and can even recognize some emotions expressed in music in early childhood, the question arises as to *how* we identify emotion in music.

One possibility is that listeners may be drawing on similarities to emotional cues that are not specific to music. For example, similarities have been drawn between the cues we use to decode emotions in music and the emotional cues we use to interpret emotion in human vocal communication. Specifically, *speech rate* (which is similar to *tempo* in music) and *voice intensity* in speech (similar to loudness or *sound intensity* in music) both increase when expressing happiness and anger, and decrease when conveying sadness and tenderness in speech or music (Juslin & Laukka, 2003). These expressive features of speech serve as basic psychophysical cues that convey the speaker's intended emotions with some level of success, even for listeners who are unfamiliar with the language being spoken (e.g., Scherer, Banse, & Wallbott, 2001) and for infants and very young children (Fernald, 1989; see also chapter 8 in this volume).

Thus, one way listeners may recognize emotions in music is through detecting general similarities in speech prosody. ***Prosody*** refers to the 'musical' qualities of speech – such as variations in pitch, pace, rhythm and stress patterns, and loudness – that convey meaning to the listener. For instance, happiness/joy is conveyed through a fast pace of speaking, a generally higher range of pitches, and an energetic bright tone quality in the voice; in contrast, sadness is conveyed through a slower pace of speaking, a lower range of pitches, and softer or less energetic quality in the voice (Juslin & Laukka, 2003). Parallels can easily be drawn to the expression of 'happiness' in music through faster tempo, higher register, and greater loudness or intensity, or the expression of 'sadness' in music through slower tempo, lower register, and softer tone quality (Gabrielsson & Lindström, 2010). Thus, although representing different channels of communication, vocal expression and music performance seem to rely on the same codes or acoustic cues to convey emotion (Juslin & Laukka, 2003). This may explain why some studies show that musically trained individuals are better at decoding speakers' emotions in everyday speech (Lima & Castro, 2011b) and even in a foreign language (Thompson, Schellenberg, & Husain, 2004) than individuals without musical training.

An intriguing study by Curtis and Bharucha (2010) further suggests 'that human vocal expression and music share an acoustic code for communicating sadness' (p. 335). The researchers asked nine female actors to say two-syllable phrases (e.g., 'Let's go,' 'Okay') to express four emotions (happiness, sadness, anger, and pleasantness) to fit an emotional scenario. They found that the actors tended to speak the two syllables in speech tones that resemble a *falling minor third interval* (-300 cents) when conveying 'sadness.' The speech samples can be heard at http://curtislab.info/emotion2010.html. When listening to the 'descending minor third' samples for 'sadness,' we can easily imagine how they map onto the falling minor third in music – such as in the interval sung to the lyric 'Hey Jude' in the opening of the song. As the actors were free to perform the phrases in their own way to convey each emotion, it is interesting to note the parallel between the expression of sadness through music in minor keys and the minor third interval in melodies that commonly express sadness in Western music (see Cooke, 1959).

In the next part of the study, Curtis and Bharucha (2010) played the actors' recorded speech phrases to listeners, and found that the interval of a falling minor third was also the best predictor for listeners' *perception* of sadness in speech. A few other intervals were also commonly used to convey an emotion, such as an ascending minor second or ascending perfect fifth to convey 'anger.' The researchers stated that these findings 'suggest that the acoustic code for communicating emotion across the domains of music and speech is more similar than previously thought' (p. 346).

Parallels between speech and music have also been examined in cross-cultural studies. For instance, fluctuations in pitch in the speech prosody of Indian Tamil speakers conveying positive/excited and negative/subdued emotions resemble the melodic interval patterns in classical South Indian music expressing the same emotions (Bowling, Sundararajan, Han, &

Purves, 2012). In general, joyful speech and music tend to be marked by larger intervals than sad speech and melodies. This pattern has been found in many styles of music – including Western classical music, South Indian 'carnatic' music, and Finnish folk music – and may reflect the biological purposes of vocalization (Bowling, 2013). For more on the topic of culture and emotion in music in familiar and unfamiliar styles, especially as it pertains to recognizing or perceiving emotions in music, readers are referred to chapter 15.

Felt emotion in music: Emotions induced by music

Thus far, we have discussed how music may express emotion to listeners. But can music also elicit emotions in the listener? And if music is able to bring about emotions, do listeners experience the same emotions – or do these 'felt emotions' vary widely between individuals? This fundamental question links with our discussion in the previous chapter on the *emotivist–cognitivist problem* (see also Harré, 1997). The idea that music expresses or represents emotions, but does not elicit real emotions is consistent with the ***cognitivist*** position, whereas the view that music can also stir up felt emotions in listeners has been characterized as the ***emotivist*** position (see Kivy, 1990), as discussed in more detail in chapter 13.

Self-reports of physiological responses

An early investigation by John Sloboda (1991) examined the relationship between structural and dynamic properties of music and reports of physical reactions they evoked in listeners. Participants were asked to recall any pieces of music to which they remembered having experienced a physical-emotional reaction (such as a bodily sensation), and to identify specific parts of the music that provoked these reactions. Sloboda then tracked down and painstakingly analyzed the musical transcriptions of every piece of music the participants had spontaneously recalled, identified the specific location and features of the music that provoked the emotional reaction, and drew links between the specific features in the music and the recalled physical-emotional responses.

Three categories of physical-emotional responses to music were identified: 'Tears' (e.g., lump in the throat, crying) were most often provoked by melodic appoggiaturas, in which a tone in the melody provides temporary dissonance and then resolves to a harmonious tone. For example, one piece that provoked tears was the opening passage from Albinoni's *Adagio for Strings*, which includes a series of appoggiaturas that shift from dissonance to consonance. 'Shivers' (e.g., chills down the spine, goose bumps) were mainly provoked by a sudden or unprepared change in harmony. 'Heart' responses (e.g., a racing heartbeat) often coincided with sudden changes in texture or dynamics or a musical event occurring earlier than anticipated. For instance, heart reactions were often provoked by the last movement of Beethoven's *Piano Concerto No. 4 in G major*, where a regular pattern of phrases is broken as one phrase suddenly comes a measure earlier than expected at bar 91, and is suddenly also much louder. Consistent with the *emotivist* view, the findings showed that listeners do report strong emotional and physical responses to music, and that the musical features that seem to trigger these responses can be identified with some degree of consensus among listeners.

Peripheral physiological responses

Many studies, such as Sloboda's research, have provided self-reports of intense responses elicited by music. But is there more objective evidence that listeners actually *experience* these emotional responses?

One phenomenon that has been shown in several studies to correspond with specific physiological changes is the 'chill' or 'frisson' response to music (e.g., Grewe, Nagel, Kopiez, & Altenmüller, 2007; Guhn, Hamm, & Zentner, 2007; Salimpoor, Benovoy, Larcher, Dagher, & Zatorre, 2011). The ***music-induced chill***, or ***frisson***, refers to 'a sudden, arousing reaction that is accompanied by goose bumps, shivers, or tingles in the spine' (Guhn et al., 2007, p. 473), similar to Sloboda's 'Shivers' categories. For instance, in Guhn and colleagues' study, reported chills coincided with particular patterns of heart rate and increases in skin conductance. An analysis of the musical passages that tended to elicit chills showed that they shared some common features: Chills tended to occur during slow movements, when a solo instrument emerged or became distinct from the accompaniment, when there was an increase in loudness or a swell in the music, or a sudden expansion of pitch range, and (as Sloboda also found) when the harmonic progression was unusual or unexpected. Participants' reports of music-induced chills have been corroborated by observed changes in skin conductance (such as in Guhn et al.'s aforementioned study) and changes in blood flow to specific regions of the brain (e.g., Blood & Zatorre, 2001).

What about more specific felt emotions? Do our bodies undergo different physiological changes when listening, for instance, to music that stirs up feelings of happiness versus sadness or fear? To address this question, Carol Krumhansl (1997) compared self-reports of felt emotion with peripheral nervous system responses to music in one of the important early studies in this area.

In Krumhansl's study, six selections of classical music were presented to participants. One group of participants was asked to indicate the degree of sadness, fear, happiness, and tension they experienced by adjusting a slider on a computer *while listening* to the music. After listening to each excerpt, they were also asked to rate how they felt while listening to the music on a set of 13 scales representing emotions such as 'afraid,' 'angry,' 'anxious,' and 'surprised.' A second group of participants listened to the same six musical excerpts while their physiological responses were recorded by polygraph. Among the measures used were heart rate, blood flow, breathing, and temperature and dampness of skin. After each excerpt, they also completed the set of 13 rating scales as the 'slider' group.

The measures of felt emotions taken during listening revealed that the six musical selections induced three distinct emotions: The Albinoni (*Adagio in G minor*) and Barber (*Adagio for Strings*) excerpts yielded high ratings for sadness; the Holst ('Mars' from *The Planets*) and Mussorgsky (*Night on Bald Mountain*) yielded high ratings for fear, and the Vivaldi ('Spring' from *The Four Seasons*) and Alfvén (*Midsommarvarka*) yielded high ratings for happiness. Further, the physiological readings also systematically differed for these pairs of excerpts, indicating that emotion-specific physiological responses correspond with the particular emotion in the music. In general, slider movements along the 'sadness' dimension corresponded with changes in cardiac and electrodermal systems, the 'fear' dimension with cardiovascular changes, and the 'happiness' dimension with changes in respiration measures (p. 349).

Krumhansl's (1997) findings show that 'not only do listeners verbally report emotional responses to music with considerable consistency, music also produces physiological changes that correspond with the type of musical emotion' (pp. 350–351). These findings have generally been corroborated in subsequent research (e.g., Gomez & Danuser, 2007; Lundqvist, Carlsson, Hilmersson, & Juslin, 2009; though see also Blood & Zatorre, 2001). Krumhansl concluded that the results of her study generally support the *emotivist* view of musical emotions. The findings also point to the dynamic nature of musical emotions. Few correlations were found between averaged physiological measures and the ratings on the 13 self-report scales taken *after* each excerpt was played, suggesting that musical emotion may not be well captured in a single 'snapshot' reading after the fact. The dynamic and temporal character of emotional responses to music may be more faithfully represented in measures taken across time as the music unfolds.

Prevalence and scope of felt emotions

Listeners report a wide range of emotions aroused by music, including feeling *happiness–joy*, *pleasure–enjoyment*, *sadness–melancholy*, *nostalgia–longing*, *love–tenderness*, *interest–expectancy*, and *pride–confidence* evoked by music (see Juslin, 2013b). It is difficult to ascertain how frequently music arouses emotions in the listener; by some estimates, music induces emotions in about 55–65% of listening episodes (Juslin, 2013a), though frequency and intensity varies widely between individuals. For example, individuals who score high on scales measuring 'empathy' (the ability to understand thoughts and feelings of others) and 'openness to experience' (encompassing qualities such as aesthetic sensitivity, attentiveness to inner emotional states, vivid imagination, and curiosity) report feeling more intense emotional responses when listening to sad music (Vuoskoski, Thompson, McIlwain, & Eerola, 2012).

Musically trained adults also report feeling stronger emotions such as happiness, sadness, peacefulness, and fear/threat in response to music, compared to those without musical training (Lima & Castro, 2011a). At the same time, musically trained individuals also more frequently report listening to music for cognitive or intellectual stimulation, as opposed to purely emotional reasons (Getz, Marks, & Roy, 2014), so they seem to engage with music in several different ways. Further, some neuroscientific studies have corroborated these reports, showing an increased sensitivity to expressive music performance in individuals with musical training (e.g., Chapin, Jantzen, Kelso, Steinberg, & Large, 2010), although both musically trained and untrained listeners report being emotionally engaged with music. In other words, musical training seems to alter the affective impact of expressive music on the brain.

As for mixed emotions, these are infrequently reported in response to music. For instance, mixed emotions were reported in only about 11% of music listening episodes according to a survey (Juslin, Liljeström, Laukka, Västfjäll, & Lundqvist, 2011). A study in which participants were asked to press a button every time they felt happiness or sadness showed that they often pressed both buttons at the same time while they listened to music with conflicting affective cues (e.g., major key: slow tempo, minor key: fast tempo; Larsen & Stastny, 2011). This finding suggests that listeners experience mixed emotions simultaneously, even if oppositional, such as happy/sad, as opposed to alternating between feeling two or more discrete emotions. Studies have also found that listeners' mixed happy/sad responses to music with mixed affective cues (e.g., major: slow, minor: fast) tend to be stronger for perceived emotions than for felt emotions (Hunter et al., 2010).

Emotion regulation and music

Given the physiological effects of music and its ability to induce emotion and modulate arousal, it is not surprising that people often use music to regulate their emotions. ***Emotion regulation*** refers to the processes by which we manage what emotions we feel, and influence the intensity, duration, and physiological and behavioral expressions of those emotions.

A number of studies have shown that many people use music for emotion regulation throughout their lives. For instance, Saarikallio (2011) interviewed people from about 20 to 70 years of age, and found that the goals of emotion self-regulation through music remain much the same from youth through old age. These include: (1) to relax and revive oneself;

(2) maintain a positive mood; (3) distract or divert attention from negative thoughts; (4) provide a framework for the mental work involved in processing internal conflicts; (5) energize or 'psych' oneself up; (6) stir up and access a range of intense feelings; (7) discharge anger or other negative emotions; and (8) provide solace when dealing with loss or loneliness.

Examining more specific contexts for how music is used in everyday life, Hallet and Lamont (2016) found that not only do many people use music to help them achieve their goals during exercising, but it also serves as an effective pre-exercise motivator. Lamont and Heye (2010) also found that with the proliferation of portable music players, many people also frequently report listening to music while traveling. During travel, music staves off boredom and brings about positive moods, and creates an 'auditory bubble' that insulates listeners while also (paradoxically) heightening their senses to their environment.

In general, people seem motivated to shift their moods in a positive direction, although this depends on the individual's current motivational goals. For instance, Tahlier, Miron, and Rauscher (2013) asked college students to spend 10 minutes writing about either an event in their lives that made them very sad but that had since been resolved, or a very sad event that was still unresolved (i.e., that they had not come to terms with yet). When asked what music they would want to listen to immediately afterwards, those who had been reminded of a sad *unresolved* event expressed greater desire to listen to music that was happy, upbeat, exciting, and active than those in the resolved sad condition. This may be because individuals experiencing unresolved sadness are more motivated to lift their mood for the energy boost they need to resolve an issue, to induce a happier mood, or to distract themselves from the lingering feelings of sadness, compared to those who had already resolved the issue.

Rumination

One intriguing strand of research emerging in this area has examined adaptive and maladaptive forms of emotion regulation or mood management using music, especially among individuals with an inclination toward rumination. ***Rumination*** refers to the tendency to respond to negative emotional events by repeatedly focusing on the distressing feelings, consequences, and possible causes surrounding a negative emotional experience. Although it may be adaptive to 'reflect' on negative experiences to work through the distress and gain some insight, ruminators tend to dwell on the negative effects in a manner that seems to maintain their distress and negative emotions.

For instance, Chen, Zhou, and Bryant (2007) temporarily induced sadness by showing college students excerpts of sad television programs, and then asked them to select music to listen to from a website with a wide range of popular music. The participants were free to peruse and listen to any music they wished for eight minutes. Chen and colleagues found that most participants initially listened to sad music that matched the mood that had been induced, but shifted to joyful music about halfway through the time allotted for browsing. On the other hand, participants who reported a strong tendency to ruminate spent 80% of the time listening to sad music and avoided the lively/joyful music, thus reflecting a pattern of behavior that maintained the negative mood and encouraged it to linger. Another study showed that ruminators actually benefited more from listening to happy music after sadness was induced than non-ruminators (Garrido & Schubert, 2015). Thus, it appears that ruminators tend to choose music that sustains or prolongs their negative mood even though they may stand to benefit even more than the typical person from listening to happy music.

A similar pattern has been observed in individuals with depression. Whereas non-depressed individuals tend to turn to music for revitalization and inspiration, individuals with depression tend to select music that reflects their mood – including when feeling sad, depressed, or other negative emotions (Wilhelm, Gillis, Schubert, & Whittle, 2013). As listening to sad music has been found to increase the depressed mood of individuals with a tendency toward depression (Garrido & Schubert, 2015), the use of music for rumination may prolong or even worsen negative moods.

Why do people like to listen to sad music?

Despite all we have described so far about sad music and negative moods, a liking for sad music does not have to be maladaptive. Listening to sad music may only be problematic when individuals are compelled to use music to sustain or amplify negative moods in a way that hinders effective daily functioning. Indeed, many listeners report that they enjoy listening to sad music. For instance, just over half of the college students in Garrido and Schubert's (2011) study agreed or strongly agreed with the statement 'I like to listen to music which makes me feel sadness or grief' (p. 286). Listeners report that sad music does not arouse just feelings of sadness, but also more positive and complex emotions such as nostalgia, wonder, tenderness, and peacefulness (Taruffi & Koelsch, 2014; Vuoskoski et al., 2012), and many listeners report deriving rewards (such as consolation and comfort) from listening to sad music (Taruffi & Koelsch, 2014).

The question of why people enjoy listening to sad music (or partaking of a wide range of art that expresses themes of sadness or grief, such as in paintings and novels and films) is a matter of debate in several fields. When it comes to music, this so-called ***paradox of tragedy*** or ***paradox of pleasurable sadness*** has been explained in a number of ways. Schubert (1996, p. 25) proposed that when sadness or other negative emotions are aroused in an aesthetic

Figure 14.2 Listeners report that listening to sad music doesn't just induce sadness, but also evokes many positive and complex emotions, such as nostalgia, wonder, and peacefulness.

Source: Photograph used with permission © Malavika Rao and Austin Vance.

context, it activates a cognitive mechanism he refers to as a ***dissociation node***, which in turn inhibits the displeasure center of the brain. A pleasurable response thus ensues from the arousal produced by the aesthetic experience, without the accompanying displeasure. Huron (2011) explains 'pleasurable sadness' derived from music by speculating that in some individuals, music that induces sadness may trigger the release of ***prolactin***, a hormone that counteracts the effects of intense negative emotions. The listener thus derives pleasure from experiencing the consoling and soothing effects of this hormone without the pain that usually triggers its release.

Juslin's (2013a) BRECVEMA model (named for eight mechanisms by which he proposes music induces emotions: *brain stem reflexes*, *rhythmic entrainment*, *evaluative conditioning*, *contagion*, *visual imagery*, *episodic memory*, *musical expectancy*, and *aesthetic judgment*) offers another view on the paradox of tragedy that focuses on ***mixed emotions*** (pleasure and sadness). Specifically, we experience sadness (contagion) simultaneously with the perceived beauty of the music (aesthetic judgment), and this may move us because these mixed emotions reflect the mixed character of life. As philosopher Levinson (2004) has expressed, because life is inherently both sorrowful and pleasurable, music that reflects this mixed bittersweet character may be most moving to us.

More recently, Sachs, Damasio, and Habibi (2015) have offered a framework that hinges on the idea of ***homeostasis***: Central to this view is the idea that music can arouse a pleasurable response if there is an imbalance in one's internal conditions that interferes with optimal functioning, and if music can restore the balance. They propose that humans are driven by a biological need to maintain psychological and physiological balance, and that this combines with individual and situational factors that allow for sad music to regulate homeostasis. For instance, an individual who has just experienced an event that triggered emotional distress and has a highly absorptive personality (i.e., a tendency to become intensely focused) may listen to sad music to shift their focus to the music and thus regulate his or her emotions toward a more optimal state of well-being. Another individual who is highly open to experience (i.e., seeks new and diverse experiences) may be drawn to sad music in the absence of any distress, simply because it meets his or her strong need for variety and novelty, and thus may bring about greater equilibrium.

At this time, there does not yet seem to be consensus in the field on which of these views are most tenable, or how they may complement each other, and more investigations are needed to test these theories. Some studies have identified individual differences associated with the enjoyment of music expressing or evoking negative emotions (e.g., Garrido & Schubert, 2011; Vuoskoski et al., 2012), as well as situational factors and transient factors such as mood (Sachs et al., 2015; Taruffi & Koelsch, 2014), so it is possible that different theories may apply to different individuals, contexts, and perhaps also to different genres of music.

Neural bases of felt emotions to music

Krumhansl's study showed that listening to music can induce emotions that correspond to specific physiological changes in the body. What happens in the brain as a mood or emotion state is elicited by music? As mentioned earlier, we have some physical evidence that peak experiences in music – the experience of 'chills' – are rooted in basic neural reward systems. Many brain areas contribute to this system and respond to music that induces chills in listeners (Blood & Zatorre, 2001). But perhaps the most important component is the production of ***dopamine***, a neurotransmitter that (among other things) communicates reward in the brain. When you feel the surge of positive feelings after winning the first prize in a contest – that's dopamine. The dopaminergic system helps us learn, thrive, and stay alive. Dopamine also contributes to the experience of chills in music (Salimpoor et al., 2011).

Salimpoor and colleagues (2011) demonstrated this important finding in an experiment that harnessed critical strengths of behavioral research, PET, and fMRI. Behaviorally, the researchers addressed a difficult methodological issue: The fact that music associated with 'peak' experiences varies widely from person to person. Through an exhaustive screening procedure that started with 217 individuals, the researchers winnowed down to a sample of eight whose musical tastes were counterbalanced. Specifically, musical excerpts that led to peak emotional experiences for some of the sample were considered neutral for others, such that each excerpt functioned both as a 'control' and as an 'experimental' condition, depending on the participant (a strategy that was introduced by Blood & Zatorre, 2001). The researchers were thus able to measure highly personal emotional responses to music, while maintaining experimental control over musical features that may lead to those responses.

The experimenters went on to use a feature of PET in order to identify neural sites in the ***basal ganglia*** (see chapter 4) that produce dopamine. In chapter 4, we learned that PET typically measures the activity of a tracer that binds to glucose. Salimpoor and colleagues (2011) used a different kind of tracer that binds to dopamine receptors, and thus showed that these receptors were more active during music that produced peak experiences. A subsequent fMRI study looked at changes in neural activity at these sites, to show that the timing of dopamine activation was linked to specific moments of peak activation – and even in anticipation of these experiences! In short, even though music expresses emotions often in abstract ways (see chapter 13 for further discussion), it stimulates emotional responses that are linked to basic survival mechanisms.

Other research using fMRI has focused on specific musical features that arouse an emotional reaction. For instance, evidence from an fMRI study suggests that ***consonant*** musical intervals (which are pleasant and harmonious to the ear) stimulate a region of the *orbitofrontal cortex* (interior relative to the region shown in Figure 4.2), which is associated with reward and reinforcement (Blood, Zatorre, Bermudez, & Evans, 1999; for a similar result with longer musical excerpts, see Mitterschiffthaler, Fu, Dalton, Andrew, & Williams, 2007). By contrast, ***dissonant*** intervals increase activity in the *parahippocampal gyrus*, a region that has intricate connections to the amygdala, the brain's 'warning center.' Further evidence of the role of the amygdala as a warning center extending to music comes from a study showing that removal of the amygdala (to prevent seizure) leads to a reduced ability to recognize 'scary' music (Gosselin et al., 2005).

Application: The emotional power of music in film

We conclude this chapter by applying the topic of emotion in music to a real-world context that has general relevance in our lives: the role of music in film. In the context of film music, not only is an external idea supplied as in the imaginative title or synopsis in program music, but concrete visual images accompany the music within the context of a storyline and fictional characters. Does music still carry emotional weight and meaning of its own, or is the emotional meaning driven primarily by the vivid moving image?

Music has always played an important role in motion pictures. Even in the days of so-called 'silent' film, music supplied by live musicians masked the sound of noisy projectors and restless audiences, and broadly underscored the general mood and actions shown on screen. Since the introduction of sound films in the 1920s, the soundtrack has played an increasingly important role in contributing to the immersive quality of film. Film sound is not contained in a flat, two-dimensional screen of a prescribed size, as are film images

(Chion, 1994). Thus, sound adds 'a third dimension' to the film experience, as the audience is immersed in an 'envelope of sonority.' The viewers gaze at a rectangular screen to which most conscious attention is directed, but are surrounded on all sides by sound. While watching a character fleeing from imminent threat during a heavy downpour, we hear the sinister heavy breathing and accelerating footsteps following right behind us, as the roar of rain showers surround us on all sides. Sound thus places the viewer *in the center* of the experience, rather than as a spectator (Doane, 1980).

Film music has the power to further immerse the viewer by heightening emotions. For instance, in an early study focusing on how music *induces* emotion during film viewing, Thayer and Levenson (1983) found that physiological responses to a graphic film about industrial accidents were significantly stronger when the film was paired with a 'horror' music track than with more neutral 'documentary' music. A more recent brain imaging study also showed that when either joyful or fearful music was paired with an emotionally neutral film clip, the combination of film *and* music together evoked stronger signal changes in the amygdala (which plays a key role in processing the emotional significance of stimuli) than when the film or music was presented alone (Eldar, Ganor, Admon, Bleich, & Hendler, 2007).

Music that *expresses* emotion has also been shown to intensify the perception of the emotional content of images. For example, Bolivar, Cohen, and Fentress (1994) paired 'aggressive' or 'friendly' music with filmed interactions of (real) wolves that were either friendly or aggressive. Overall, the aggressive interactions were rated by viewers as significantly more aggressive than the friendly interactions. However, within the aggressive and friendly interactions, the music either increased or reduced the perceived degree of aggressiveness or friendliness in ways that were congruent with the emotion expressed by the music.

Of course, the score and film images do not always mirror each other with respect to emotion or mood. In certain cases, the music may be 'anempathetic' – or seemingly indifferent – to the emotion or mood of the scene. After the horrific murder takes place in the classic shower scene in Hitchcock's *Psycho* (1960), the sound of the running shower continues 'as if nothing had happened' (Chion, 1994, p. 9). A memorable musical example occurs in another Hitchcock film, *The Birds* (1963), in which a woman sits on a bench in a school playground, smoking a cigarette. Crows ominously gather behind her, until she finally turns to discover thick flocks of birds covering the playground and rooftops, and flees in fright. Meanwhile, the sound of children singing a nursery rhyme inside the schoolhouse carries on relentlessly throughout this chilling scene. As Chion has observed, anempathetic music can create the effect 'not of freezing emotion but rather intensifying it' (1994, p. 8).

Boltz (2004) proposed that information from images and music that are juxtaposed in mood in this way may be encoded separately, as opposed to being processed in an integrated fashion. She found that memory for the details about a film scene and the accompanying music were significantly poorer when the images and music were mood-incongruent than when they were matched for mood. Further, when viewers were instructed to focus only on the images or the music, incidental learning occurred in the other mode, but only for mood-congruent films. It is possible that mood-congruent information provided in more than one sensory modality is more likely to be integrated into memory as a unified whole than cross-modal information that is mood-incongruent.

Diegetic and nondiegetic music

A standard method used in film music research is to present the same film clips with different music tracks to examine how viewers' responses differ with the music (see Tan, 2017a).

However, manipulations of the same soundtrack can also have powerful effects. For instance, Tan, Spackman, and Wakefield (2017) examined the effects of presenting the *same* piece of music to suggest diegetic or nondiegetic music. The term *diegesis* refers to ‘all that belongs … to the world supposed or proposed by the film’s fiction’ (Souriau, quoted in Gorbman, 1987). Thus a song playing over a juke-box during a bar-room brawl is supposed to be ***diegetic music***, as it seems to exist within the fictional world inhabited by the characters. On the other hand, a fast music track mirroring a car chase is usually assumed to be ***nondiegetic music***.

Tan, Spackman, and Wakefield (2017) selected a scene from Spielberg’s *Minority Report* (2002), showing a man and woman making their way hurriedly through a shopping mall while being pursued by police (see Figure 14.3). The scenario and the relationship between the two characters are rather open to interpretation for those who have not seen the film. In the original Spielberg film version, the music (an instrumental version of Henry Mancini’s song ‘Moon River’) sounds faint and distant, as if playing over loudspeakers inside the mall (*diegetic* music). The researchers obtained the ‘Moon River’ single and laid the same music track over the soft music, mixing it louder and clearer to suggest a dramatic score (a *nondiegetic* music version).

The findings showed that participants interpreted the scene differently – depending on how the ‘Moon River’ music was presented. Those who watched the *diegetic* music version (in the original *Minority Report* film) perceived the scene to be more tense and suspenseful, assumed a more hostile and antagonistic relationship between the two characters, and believed them to be more fearful and suspicious of each other and more intent to harm each other compared to those who had watched the nondiegetic version. Participants who saw the *nondiegetic* version (with the same music mixed to suggest a dramatic score, external to the world of the characters) also perceived the male as experiencing less fear, less excitement, and more romantic interest in the other character. It is possible that Mancini’s gentle ballad ‘Moon River’ was perceived as incidental when it sounded like mall music that happened to be playing within the environment of the characters, but as a commentary to the scene when presented as the dramatic score.

The calm, gentle music was mismatched with the suspenseful chase scene in *Minority Report*; presenting it diegetically may have allowed the film-maker to introduce the incongruent music track to dramatic effect in a convincing/natural way. Indeed, another study showed that participants gave higher ratings for how ‘good’ they judged an audiovisual combination to be when the music that did not match the mood or action of a scene was clearly diegetic (i.e., the source was shown on the screen) than when the source of the music was more ambiguous (Fujiyama, Ema, & Iwamiya, 2013).

Interpretations of storyline and characters

Film music can also wind up tension and heighten emotions by influencing the audience’s interpretation of the unfolding storyline, and shaping their predictions and expectations about future events (see Tan, 2017b). For instance, Vitouch (2001) showed an open-ended scene from Billy Wilder’s film *The Lost Weekend* (1945) with one of two music tracks, and asked participants to describe what the man was thinking and what would happen next. The scene shows a panoramic shot of New York City and pans over to an open window, through which a man can be seen packing a small case. Viewers who watched the scene accompanied by the (original) Rózsa score, which is a serene and lush orchestral piece, tended to write positive or ambivalent story continuations, for example: ‘The man is in a good mood, has a

Figure 14.3 Six images from *Minority Report* (DreamWorks and Twentieth Century Fox, 2002), directed by Steven Spielberg and produced by Gary Goldman and Ronald Shusett, actors are Tom Cruise and Samantha Morton. This excerpt was used in Tan and colleagues' (2017) study.

Source: Please see the reference list under the director's name for film information.

secure and well-paid job, and is just preparing a meal for his new love: candle-light dinner for two . . .' (Vitouch, 2001, p. 76, positive continuation).

However, participants who watched the same scene accompanied by an excerpt from Barber's *Adagio for Strings*, which had been found in previous research to induce 'sad' or 'melancholic' emotions, were more likely to write negative continuations, for instance: 'Why did this happen to me? How shall I go on? . . . He's just been left or a very close person has died . . .' (p. 77). As Cohen (2009) has proposed, 'the audience is constantly synthesizing a story, a *working narrative*, derived jointly from sensory information and hypotheses or expectations based on long-term memory' (p. 443), and 'cues from music as well as from the other sources of information contribute to the working narrative' (p. 444).[1]

Different music tracks may also lead to different assumptions about the details within this working narrative, such as the intentions of a character or the nature of the relationship between characters (e.g., Boltz, 2001; Bullerjahn & Güldenring, 1994), and may even affect how much an audience likes a character (Hoeckner, Wyatt, Decety, & Nusbaum, 2011). In Boltz's (2001) study, for instance, participants who saw a scene accompanied by music conveying a 'positive' mood (major key, predictable rhythm) were most likely to interpret the relationship between two characters in a scene as harmonious or romantic. Participants viewing the same scene with music conveying a 'negative' mood (minor key, irregular rhythm), on the other hand, were more likely to infer that one character would harm another. Positive music also led to more positive descriptions of a male character's traits (e.g., kind, loving, protective), whereas more negative personality descriptions were ascribed to a male character (e.g., deranged, evil, manipulative) when the same scene was accompanied by negative music.

Music may not even need to accompany a character to influence the viewers' perceptions of him or her. Tan, Spackman, and Bezdek (2007) paired film scenes that featured a solitary character (showing no strong emotion) with nondiegetic music that had been shown to convey 'happiness,' 'sadness,' 'fear,' or 'anger' in previous studies. (Still frames are shown in Figure 14.4.) The music was played either *before* a solitary character entered the scene or began *after* the character left the scene – overlapping for only a few seconds with the music to create a smooth transition, as is the practice in film editing. Viewers tended to perceive the character's emotions in ways that were generally congruent with the emotion expressed by the music track, even though the figure was mainly shown without any accompanying music (only sound effects). Tan and colleagues' study showed that music not only influences our interpretation of concurrently shown images, but may influence viewers' expectations about forthcoming scenes and reframe their impressions of images they have just viewed (see also Boltz, Schulkind, & Kantra, 1991).

Closure

Musical tension resolution in the score, as discussed in the theories of musical structure in chapter 7 and Meyer's theory in chapter 13, may also have an impact on the perception of the structure of the film. For example, a study by William Forde Thompson and colleagues showed that film excerpts are judged to have ended with greater closure if accompanied by a music track that resolves to the tonic (e.g., ending on a C chord if in the key of C) than if the music remains unresolved (Thompson, Russo, & Sinclair, 1994). Further, the effects are stronger for tonally closed music with a clear metrical structure and ending on a strong beat.

A striking example of weak versus strong closure can be found in the second and third films in the Wachowskis' *The Matrix* trilogy, scored by Don Davis. The final scene of *The Matrix Reloaded* (2003) ends with an unresolved progression of dramatic chords that leaves

Figure 14.4 Images from three films used in Tan and colleagues' (2007) study: *Swimming Pool* (Fidélité, 2003), directed by François Ozon, produced by Olivier Delbosc and Marc Missonier, actor is Charlotte Rampling; *Three Colors: Blue* (CAB Productions, 1993), directed by Krzysztof Kieslowski, produced by Marin Karmitz, actor is Juliette Binoche; and *Diva* (Les Films Galaxie, 1981), directed by Jean-Jacques Beineix, produced by Claudie Ossard, actor is Wilhelmenia Wiggins Fernandez.

Sources: Please see the reference list under the directors' names for film information. This figure appeared in *Music Perception*, *25*, p. 140, reprinted with permission of University of California Press.

the audience hanging, whereas its sequel *The Matrix Revolutions* (also released in 2003) ends with a dominant-to-tonic cadence that gives the final scene a strong sense of closure. In some films, a new piece of music begins during the last moments of the movie and proceeds seamlessly through the end credits, such as all five of the *Bourne* films (2002, 2004, 2007, 2012, 2016) so far. This gives a sense that the saga continues. There are also films in which the music and closing image end or fade away together, such as the American flag in *Saving Private Ryan* (1998), giving a very strong impression of finality. Thus the role of music on the emotional and dramatic impact of a film is evident up to its closing scenes.

As we have seen, music provides a rich conduit of emotional communication to the film audience, apart from the information contained in the images, dialogue, and other sounds. Empirical work in this area emerged in the 1980s and has seen slow but steady growth, leading to the publication of the first book to consolidate the research on the role of music in film and other media, entitled *The Psychology of Music in Multimedia* (Tan, Cohen, Lipscomb, & Kendall, 2013). Studies show that film music is not relegated to the role of simply mirroring or intensifying the effects of the screen images, but can contribute meaning of its own that advances the narrative, juxtaposes with the scene, reframes our perceptions of preceding scenes and creates expectations about upcoming scenes, and signals whether the story is resolved or remains open-ended. Although viewers often do not consciously attend to the 'unheard melodies' of cinema (Gorbman, 1987), the emotional force of the film score – in part through its power to express and to induce emotion – influences the interpretation of images and the unfolding narrative in powerful ways.

Coda

This chapter examined the emotional significance of music. From a convergence of empirical studies employing many methods including self-report, physiological, and neuroscientific methods, we have observed the profound impact of music on emotion. Even when attention is shared among several sensory modalities as in the viewing of an opera or film, the emotional power of music in its ability to represent, express, and induce emotion is clearly manifested. One important question that emerges is whether the emotional significance of music for one cultural group persists when that music is presented to a different group. More generally, one might question whether the structural conventions that have emerged in Western music are shared among other cultures. In the next chapter, we address the issue of culture more directly by surveying the case for and against 'musical universals,' and by reviewing some of the research that has aimed to understand human musicality on a global scale.

Note

1 A full explanation of other sources leading to the working narrative can be found in Annabel Cohen's discussion of her influential *congruence-associationist model* (Cohen, 2013; Marshall & Cohen, 1988), which she describes in person at http://global.oup.com/booksites/content/9780199608157/audiovisual/ch02/.

15 Culture and music

The Beatles, and in particular guitarist George Harrison, are known for introducing many Westerners to the sounds of Indian classical music. Harrison's interest in Indian music began to flourish while filming *HELP!* In the film, Indian music and culture were used to humorous effect in a way that unfortunately reflected that time period's lack of cultural sensitivity. Nevertheless, this encounter piqued Harrison's curiosity about Indian tonal structures. Later, Harrison studied the *sitar* (a plucked string instrument) under the virtuoso performer and composer Ravi Shankar, and used the instrument in several well-known Beatles songs (for engaging anecdotes, see Harrison, 1980). One song, 'Within You Without You' from *Sgt. Pepper's Lonely Hearts Club Band* (1967), features elements of the Indian raga, which will be discussed further in this chapter. Importantly, these songs did not employ Indian music in the stereotypical and parodied fashion that was common in Western forms of entertainment at the time. Indeed, Harrison and his bandmates developed a unique sensitivity and respect for Indian classical music that played a significant role in broadening the musical palate of Western audiences.

Two points from our opening example serve as the motivation for our final chapter. First, this example shows how exposure to music from other cultures can lead to a deep appreciation of that music, even if it is quite different from the music with which one is familiar. This may be more significant than it first seems. Consider the case of language. Harrison may not have felt similarly captivated by hearing the players speak an unfamiliar language like Punjabi. There is something about music that may be inherently universal in its appeal and in our ability to comprehend it. The second point is more cautionary in spirit. Historically, non-Western music was often referred to as 'primitive,' a value-laden term that suggests a progression of musical sophistication with Western music at the pinnacle. Just as humor may cause hurt when perpetuating broad cultural stereotypes – as may have been the case in *HELP!* (and many films of its day) – researchers can unintentionally cause harm by speculating too easily about hierarchies of complexity or value in music. Today, most researchers who study different cultures try to avoid making value judgments regarding any particular kind of music.

There is a reason why we positioned this chapter last; everything we have discussed so far can be re-examined through a cross-cultural perspective. The dominant perspective in psychology is that of ***cultural universalism***. Psychologists are typically (though not always) interested in learning about features of the human mind that are not influenced by cultural constraints. In contrast, musicologists often focus more on cultural specificity (Carterette & Kendall, 1999). Cross-cultural music psychology tries to bridge this gap by addressing the relative contributions of what is culturally universal or specific to the perception and performance of music. This perspective reflects an interest in understanding mental ***processes***, which are functions present in the mind of any human from any culture. In contrast, music-specific ***properties*** are musical features that reflect a specific culture (Patel & Demorest, 2013). For example, when George Harrison was first exposed to Indian classical music, universal processes he shared with the Indian composers and musicians facilitated Harrison's cross-cultural appreciation of the unfamiliar properties of the Indian music.

It is important to note that one rarely, if ever, finds a thinker so extreme as to entirely dismiss the contribution of culture. However, it is similarly shortsighted to assume that *every* aspect of musical experience is culture-specific, as some are clearly based on universal physiology (e.g., the high-frequency recording of a dog whistle on *Sgt. Pepper's Lonely Hearts Club Band* was recorded by Paul McCartney so that his dog, but not human listeners, would be entertained). It is also erroneous to assume that anything physiologically based is therefore universal, as almost every aspect of human physiology is subject to environmental influences and neural plasticity (see chapter 4).

The challenges of cross-cultural research

People are often surprised by the dearth of scientific cross-cultural research on music. Why would a field lack research in an area of such self-evident importance? One reason may stem from researchers' awareness of their own limitations. Most contemporary music psychologists are Westerners, steeped in the European tradition of (primarily classical) music. When exploring a new topic, it makes sense to begin with a context that complements one's expertise.

A second reason may stem from the practical and methodological complexities of conducting cross-cultural research. The ideal cross-cultural study is what Patel and Demorest (2013) call a ***fully comparative*** study: one in which cultural differences are defined with respect to participants (two groups reflecting different cultural heritages) as well as the cultural basis for musical stimuli. As a result, one can look at a complete set of pairs between the culture of the listener and the culture of the stimulus (music). Ideally, each group would be completely unfamiliar with the music associated with the other group.

Unfortunately, it is difficult to find research participants whose musical experience reflects a solely non-Western culture. For instance, the study of meter perception by Hannon and Trehub (2005a), discussed in chapters 6 and 8, used a fully comparative design. The authors recruited participants from the Toronto area who had grown up in the Balkans. This group's ability to detect rhythmic changes to simple and complex meters (the latter being common in Balkan, but not in North American, music) was compared to that of a group that was enculturated in North America. Although this design was fully comparative, the groups were not perfectly matched with respect to knowledge of the other culture's music: The Westerners did not know Balkan music, whereas Balkan participants were effectively bicultural (cf. Wong, Roy, & Margulis, 2009). Another option would be to recruit in the country and region of the target population, as did Krumhansl and colleagues (2000) in their study of Sami *Yoiks* (see chapter 7). Even doing this, however, does not ensure isolation from Western music.

Other intrepid researchers have traveled to remote villages and spent considerable time winning over the residents in order to collect data. Examples include studies we discuss later in this chapter: Fritz and colleagues' (2009) research involving the Mafa, living in a remote mountain region in Cameroon, and a more recent study of Pygmies in the Congolese rainforest (Egermann, Fernando, Cheun, & McAdams, 2015). In both cases, the investigations focused on groups with no prior exposure to Western music. However, such studies are rare, and the potential to find groups who are not exposed to Western music may become more elusive with increasing globalization. For better or worse, music is becoming more homogeneous and some musical cultures are vanishing, probably more rapidly than are languages. Accordingly, the scope of study in ethnomusicology has changed. Whereas before the 1970s, ethnomusicologists often neglected the study of music combining the styles of two or more cultural sources, there is now more fieldwork on music that reflects a cultural mix of musical styles (Nettl, 2001).

There is also the issue of cultural sensitivity. It is rare to encounter research on music cognition that adopts an anthropological perspective (notwithstanding the subfield of cognitive ethnomusicology). One reason for this omission may arise from the respect researchers have for the inherent value of each culture's music. To many, 'universalist' perspectives are ultimately Westernized perspectives, and theories that couch Western notions as 'universal' implicitly position a certain cultural context (that of Western Europe) at the pinnacle. As ethnomusicologist Bruno Nettl (2001) has expressed it:

> We should . . . eliminate the notion that all musics pass through a set of stages, and that we can explain the variety of world musics by suggesting that we are observing each of them at a different stage of the same development . . . . It is not inevitable (or even likely) that a non-Western music would gradually change to become like Western music.
> (p. 13)

This concern is not mere fantasy, of course, given that explicit claims to this effect have been made periodically throughout history. Bearing all this in mind, we proceed to discuss how research has addressed some critical questions concerning cross-cultural music psychology. This chapter begins with a review of research concerning music perception abilities among nonhuman animals. This research seeks to find a kind of biological baseline for music perception, against which any cultural tendency may be compared. We then go on to consider cross-cultural research from several fundamental topics in this book: musical pitch, musical time, performance, and emotional communication.

How musical are nonhuman animals?

It may be surprising to see a chapter on culture beginning with a discussion of research that does not focus on humans. However, one can argue that some of the strongest evidence for or against universalism should come from research comparing humans to nonhuman animals (henceforth simply 'animals' for brevity). Based on the logic of evolutionary biology, one might consider music-like traits that we share with closely related animals to constitute core universal musical characteristics that have developed over time as a result of evolutionary adaptations.

There are two major limitations in studying the musical capacities of animals. First, because they don't talk, it is difficult to elicit the behaviors from animals that are under investigation. Researchers typically use techniques similar to those used with infants (see chapter 8), which can involve extensive training. Second, because music (at least of the human variety) is not an intrinsic part of the animals' culture, these experiments require them to engage in unnatural behaviors. A similar issue is associated with the study of speech perception in animals. Some researchers dismiss findings related to animal speech perception based on this problem. For instance, Trout (2001) suggests that training animals to perceive speech is akin to seeing whether humans can fly if we throw them off a rooftop!

Primates

We start our discussion with species that are most closely related to humans: monkeys and apes. Researchers have assessed whether these primates perceive three basic musical features similarly to humans: octave equivalence, consonance versus dissonance, and beat perception.

One basic feature is ***octave equivalence***. This construct, which is discussed extensively in chapter 5, is based on the fact that pitches separated by an octave (a 2:1 frequency ratio) are heard as belonging to the same pitch category, even though their physical frequencies differ. As discussed in chapter 5, octave equivalence plays a pivotal role in distinguishing musical pitch perception from the perception of pitch in other domains. If nonhuman animals do not hear octave equivalence, it is unlikely that their experience of musical pitch resembles that of humans.

Interestingly, many studies have not found evidence for octave equivalence in monkeys (for a review, see McDermott & Hauser, 2005). One exception is a study by Wright, Rivera, Hulse, Shyan, and Neiworth (2000), who conducted a series of experiments in which rhesus monkeys learned through reinforcement to distinguish various auditory stimuli, some of which were melodies such as ‘Happy Birthday’ and ‘Old MacDonald’. When presented with tonal melodies separated by an octave, monkeys behaved as if no change had occurred – thereby showing octave equivalence. Importantly, this effect only worked for tonal melodies and only when constituent tones were complex tones rather than pure tones (see chapter 2). This finding is impressive taken on its own, but does not sit well with many other findings according to research reviewed by McDermott and Hauser (2005), which generally suggests that animals (including primates) experience music as a set of fixed pitch classes (i.e., in an ‘absolute pitch’ sense) rather than as pitch relationships (‘relative pitch’). Thus one must say there is ‘mixed’ support for the idea that we inherited octave equivalence from our ancestors.

Other studies have examined whether nonhuman primates are sensitive to another basic feature of music: ***consonance*** or ***dissonance*** of musical chords (i.e., whether pitches sound pleasant or unpleasant together; see chapter 3 for further discussion). McDermott and Hauser (2004) placed cotton-top tamarins in a V-shaped cage that was configured so that a consonant chord sequence would sound if the monkeys traveled down one leg, whereas a dissonant sequence would sound if they traveled down the other leg (see Figure 15.1). The tamarins exhibited no preference for one sequence or the other; they spent similar amounts of time in both sections. In fact, a similar subsequent study suggests that tamarins prefer complete silence to any kind of music (McDermott & Hauser, 2007). By contrast, humans placed in a similar context (e.g., a room configured so that consonant chords are played when walking to one side, but dissonant chords are played when walking to the other) show the kind of specific preferences one would expect, preferring consonance to dissonance and music to silence (McDermott & Hauser, 2004, 2007).

What about primates that are more closely related to humans? In fact, there is some evidence (albeit from a single individual) that chimpanzees may prefer consonance over dissonance. In this study, an infant chimp who was lying supine would pull a cord to continue a musical stimulus that was playing, or stop pulling so that the music would switch to something different (Sugimoto et al., 2010). The chimp pulled the cord more often when hearing consonant music, thus displaying a preference for sustaining this music rather than having it switch to dissonant music. The implication of Sugimoto and colleagues (2010) is that a preference for consonance over dissonance should be a universal property of human music perception. This may not be the case, however, as we will see later.

Another basic feature of music is that it communicates a regular beat to which humans spontaneously synchronize (see chapter 6 for more discussion). Perhaps the best evidence for this comes (again) from a single case, a bonobo (Large & Gray, 2015). Significantly, this case arose when researchers saw the bonobo *spontaneously* responding in rhythmic ways to

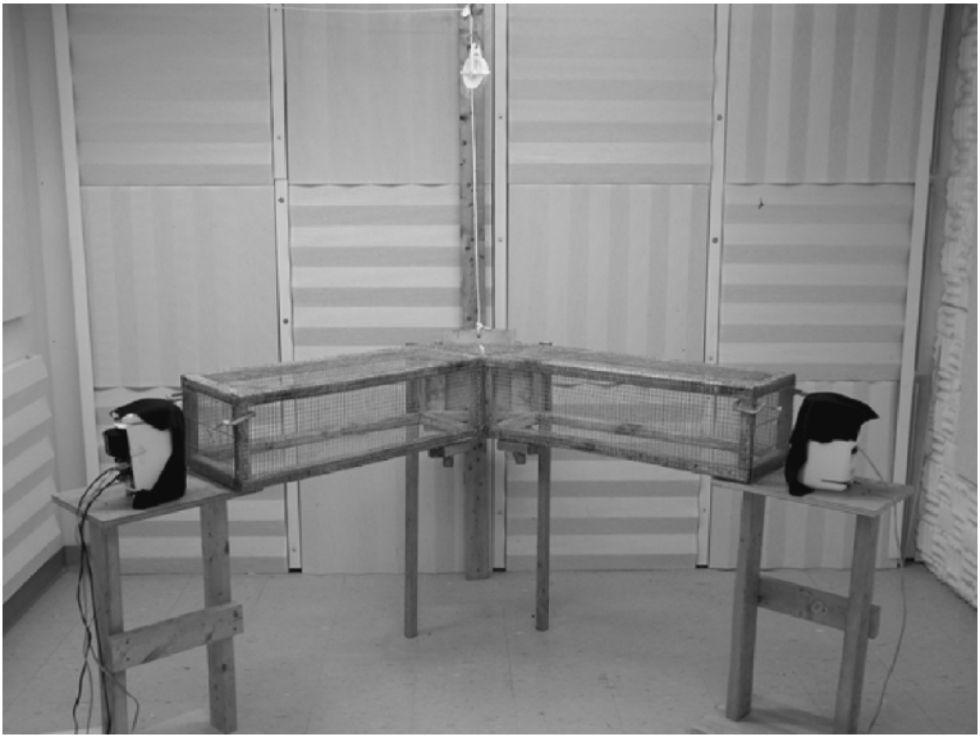


Figure 15.1 Apparatus used by McDermott and Hauser.

Source: McDermott and Hauser (2004, Figure 1, p. B13), reprinted with permission from Elsevier.

human drumming. The bonobo exhibited synchronization to an isochronous rhythm (i.e., a sequence of identical time intervals; see chapter 6) along with two critical characteristics seen in humans: flexibility across tempos and a tendency to anticipate the beat. A critical feature of this design was that the researchers measured the bonobo's spontaneous tempo (see chapter 6). Synchronization was most accurate when the isochronous rhythm was near this tempo (which again is comparable to what humans do). Other studies on primate synchronization, including chimpanzees (Hattori, Tomonaga, & Matsuzawa, 2013) and rhesus monkeys (Honing, Merchant, Háden, Prado, & Bartolo, 2012; Zarco, Merchant, Prado, & Mendez, 2009), in contrast, have found that these animals react to rather than anticipate onsets in a rhythm. These results may differ from those of Large and Gray's (2015) bonobo study because the researchers did not assess spontaneous tempo, or because of individual differences within species (i.e., spontaneous rhythmic engagement).

Birds

In contrast to other nonhuman animals, primates are most closely related to humans with respect to biology. However, birds appear to be most closely related to humans with respect to the use of music. Birds, of course, 'sing,' and many composers have based sections of pieces on the 'music' of birdsong, including Beethoven, Respighi, and Messiaen. One important limitation of primates is that their undescended larynx severely limits their vocal learning capacity because this restricts their ability to articulate (Mithen, 2006), and vocal learning may be an important component of music learning in general (Patel, 2008). Indeed, the capacity to acquire a broad repertoire of auditory sequences for production and perception, and the capacity to create novel combinations of these sequences, suggests an interesting commonality between humans and birds. But there is a limitation. Evolutionarily, such shared traits are thought to be ***analogies***, similar traits that emerge in unrelated species due to environmental demands, rather than ***homologies***, traits that are inherited from an immediate ancestor.

So how musical are birds? Birds, of course, like other animals, do 'sing.' However, it is important to include quotation marks here, because the direct comparability of birdsong to human song has been questioned. McDermott and Hauser (2005) argue that animal 'songs' (including birds, humpback whales, and mice) are far more inflexible, both with respect to structure and function, than human musical song. The structure of birdsong, for instance, is less varied than human music (at least with respect to our present abilities to define and analyze musical variety). Functionally, songs among animals seem to have very specific roles and are typically used either to attract mates or to protect one's own domain. Certainly, human music can take these roles, but we often create music or listen to music simply because it is pleasing. The musical babbling and spontaneous singing of infants discussed in chapter 9, as far as we know, does not always serve a practical purpose. In any case, one should not too casually assume that references to 'songs' among nonhuman animals constitute strong support for musicality in these species.

The way in which birds perceive pitch may also differ substantially from humans. It has been known for some time that the perception of pitch among birds is more constrained by absolute pitch than it is for humans. Whereas humans (adults and children; see chapters 5 and 8) easily recognize familiar melodies when they are transposed to a different key, birds respond to transpositions as if they are entirely different from the original melody (Hulse, Cynx, & Humpal, 1984; Hoeschele, Weisman, Guillette, Hahn, & Sturdy, 2013). As such, unlike rhesus monkeys, there is still no evidence that birds perceive pitches separated by octaves as equivalent. Furthermore, newer research calls into question whether birds perceive melodies based on the fundamental frequencies of their constituent tones. If so, birds may not really perceive pitch as humans do. A recent experiment with a sample of European starlings (Bregman, Patel, & Gentner, 2016) suggested that their recognition of auditory patterns was based on pitch or variations in the shape of each tone's power spectrum (see chapter 2). For humans, these variations would be heard as subtle difference in timbre.

So, as far as pitch is concerned, the musical capacity of birds may differ from humans, despite the apparent musical quality of birdsong. Interestingly, there may be more commonality across birds and humans when it comes to perceiving a beat. It has been demonstrated (as seen in examples on YouTube and elsewhere) that birds who engage in vocal learning can also dance to music. For instance, Patel and colleagues have shown that the sulphur-crested cockatoo called Snowball can synchronize to the beat of music, even adjusting movements (such as bobbing its head and kicking its feet) to match music played at different tempos (Patel, Iversen, Bregman, & Schultz, 2009). Similar behaviors have been studied in other parrot species, such as budgerigars trained to peck a key in time with an audiovisual pulse across a range of tempos (Hasegawa, Okanoya, Hasegawa, & Seki, 2011). Beyond birds, there is a report of a sea lion who demonstrated head-bobbing to a regular rhythmic stimulus (Cook, Rouse, Wilson, & Reichmuth, 2013), suggesting that entrainment of movement to an external rhythm may not be limited to avian species with the capacity for vocal mimicry.

Having considered the basic humanness of musical capacities (by comparing them with animals), we now consider some critical questions concerning whether musical capacities are culturally universal within humans.

Do listeners 're-tune' to pitch patterns of another culture?

One of the notable differences that can be observed across musical traditions is in the expression of pitch. For instance, Western orchestras take great pains to tune their instruments as closely as possible to a standard pitch. In comparison, pitched instruments of the Indonesian *gamelan* percussion orchestra are deliberately tuned slightly apart, such that instruments classified as *pengisep* are tuned slightly higher than those classified as *pengumbang*. Therefore, when the same note is struck on *pengisep* and *pengumbang* instruments at the same time, it produces *penyorog*, which refers to the vibrant beating or 'shimmering' quality between the tones of different instruments (Hood, 1966), a desirable effect that *gamelan* musicians strive to achieve.

In much of Western music, each pitch is typically produced as a discrete unit, with abrupt changes between pitches. By contrast, pitches in Indian music are often allowed to waver, and a performer will often glide gradually from one pitch to the next. Western listeners can often find such wavering in pitch to be disconcerting at first. However, it is important to recognize that popular music in the West often features deviations in pitch from established 'standards,' as in the glides that are representative of blues or the use of melisma that is common among singers like Beyoncé. Clearly, there are rich variations in the use of pitch, both in formal musical systems and in expressive performances. Moreover, many cultures do not use the same set of pitches or scale structure to create melodies.

Figure 15.2 (Justus & Hutsler, 2005) illustrates the diversity in pitch systems across a small sample of cultures with rich formal musical traditions. Each row in this figure includes the same set of light gray bars spaced in semitones used in Western music, corresponding to the black and white keys of the keyboard depicted below. These gray bars serve as reference points against which to compare pitches used in other cultures, which are shown as dark black bars. For most systems, the first row shows a large set of pitches from which scales can be constructed (Indonesia is the exception). These sets, when present, are generally referred to as ***tonal material***, following Dowling (1978). Note that the spacing of tonal materials (called *lü* in China) varies. In most cases, these pitches are spaced equally, though Indian music includes uneven spacing. Also, Indian and Arabic music both include much smaller pitch intervals than those found in Western music.

Each system shown in Figure 15.2 does use one or more formal ***scales*** based on five to seven pitch categories. The pitch categories that are used in scales vary widely across cultures. The diatonic system used in the West is not universal, although there are a couple similar cases (e.g., the *Bilāval* from India that approximates the Western major scale). It is clear from this figure that one should not draw universal conclusions about musical pitch perception based on experiments that use only the modern Western tonal system! In keeping with terminology introduced earlier, Figure 15.2 illustrates how diverse the properties of pitch systems are, even if there are universal processes that listeners use to perceive these systems.

Despite these differences in the realization of musical pitch, Justus and Hutsler (2005) describe some core pitch features that may be common across cultures. First, all musical cultures appear to incorporate octave equivalence in their tonal systems. This is often pointed to as a candidate universal (e.g., Brown & Jordania, 2013; though see also Tenzer, 2000). Another commonality among the scales shown in Figure 15.2 (though not always the tonal material) is that pitches are generally spaced *asymmetrically* (with the possible exception of the *Maqām Rāst* in Arabic music). We highlighted the importance of this property for Western scales in chapter 5. Based on the small sample shown in Figure 15.2, it is possible that the highly frequent (if not ubiquitous) property of asymmetry may reflect some underlying universal process. There are also some commonalities of intervals present in tonal systems.

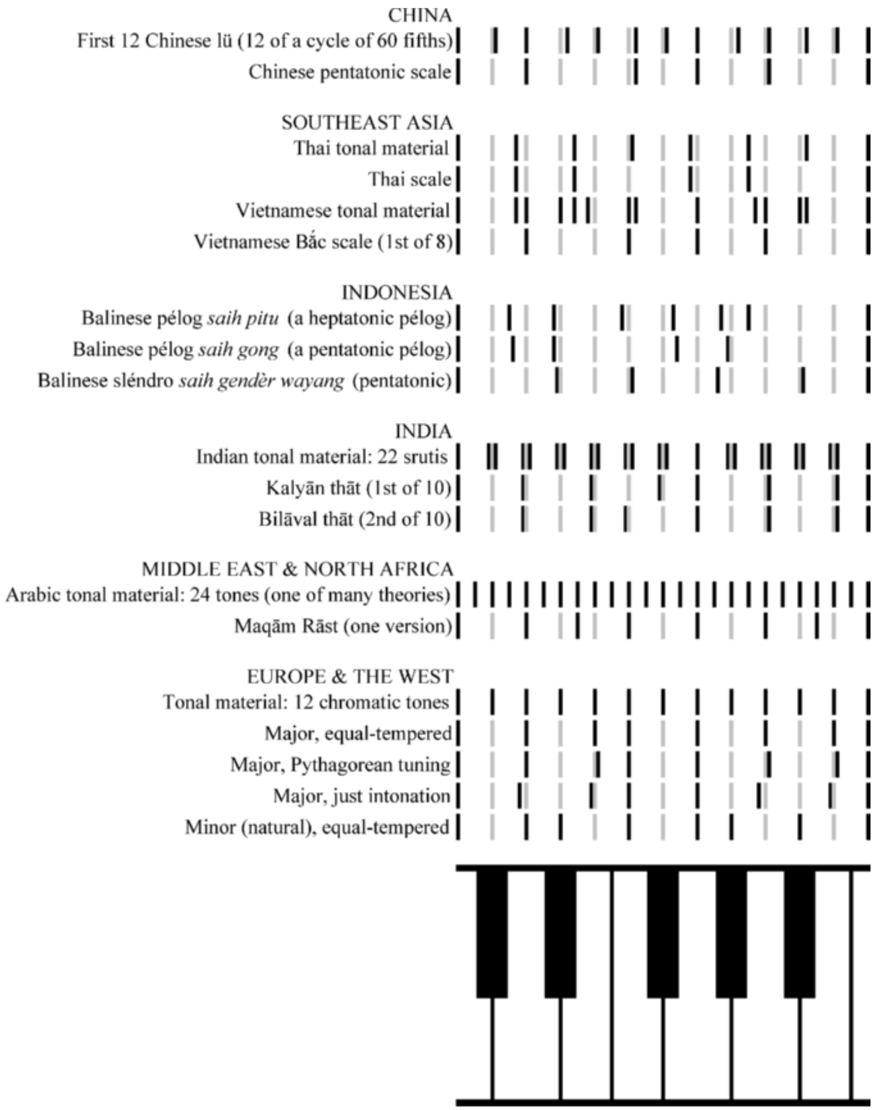


Figure 15.2 Justus and Hutsler's representation of pitch in the tonal systems of various cultures (dark bars) as compared with the 12 chromatic pitch categories of the Western equal-tempered scale (gray bars and piano keyboard).

Source: Justus and Hutsler (2005, Figure 4, p. 13), reprinted with permission from the University of California Press.

Nettl (2001) has proposed that 'the principal melodic interval in most of the world's musics is approximately a major second' (p. 9).

Next we summarize a few studies that bear on the representation of pitch structures, using research on Indian music as a model. In chapter 7, we also discussed related research on melodic expectations in Chinese and Sami *Yoik* music, and some related research has been conducted using Indonesian pitch systems (e.g., Kessler, Hansen, & Shepard, 1984; Perlman & Krumhansl, 1996), yielding analogous results to those we summarize below.

Perception of non-Western pitch structures: Indian ragas

To date, most cross-cultural research to do with pitch perception has focused on Western perceptions of Indian classical music. At the heart of Indian music is the ***raga*** (or ***rāg***), a sequence of notes taken from a scale (***thāt***) that a performer adapts while improvising melodies and rhythms. The Indian typology of emotions designates a special significance to different ragas. Each *raga* is associated with time of day or season of the year, and expresses a specific emotion. There are *ragas* for morning and evening, for spring, summer, autumn, and winter, and so on. The traditional interpretation of the affective power of certain *ragas* is still widely accepted by Indian musicologists. The notes that constitute the 'scale' of a particular raga are drawn according to a precise set of principles. Two *thāts*, illustrated in Figure 15.2, are quite similar to the Western diatonic scale. However, there are eight other *thāts* that differ more noticeably from standard Western scales. These scales create what has been called a 'circle of *thāts*.' There is a similar way to represent relationships among diatonic scales in Western music called the 'circle of fifths,' based on the fact that adjacent scales have tonics that are a perfect fifth (seven semitones) apart.

One question that has been addressed in the research, then, is whether Indian and Western listeners perceive pitch relationships in *ragas* consistent with the circle of *thāts* (Castellano, Bharucha, & Krumhansl, 1984). In Castellano and colleagues' study, Indian and Western listeners were presented first with an Indian *raga*, followed by one of several 'probe tones.' As in Krumhansl's other research (see chapter 5), listeners rated how well probe tones 'completed' the context. Indian listeners rated probe tones in ways that were consistent with Indian scale organization. The responses of Western listeners constituted a weaker approximation to the Indian scale structure. Interestingly, Western listeners did not simply assimilate pitches from Indian music to a 'Westernized' structure either. Rather, Westerners appear to have formed a partial concept of Indian music structure through the context to which they were exposed during the experiment. They did this using what can be considered a *statistical* listening strategy. That is, listeners kept track of how frequently different kinds of pitches occurred (probably without being aware), and then made judgments based on how likely pitches were within the current context. Infants may also use a similar strategy in their perception of pitch and melody (see Saffran & Griepentrog, 2001).

More recent work has addressed the perception of tonal modulations in Indian music using Carnatic music, which is more complex than the Hindustani music used by Castellano and colleagues (Raman & Dowling, 2016). In their experiment, highly skilled American and Indian listeners heard *ragas* in one ear while a constant probe tone sounded in the other ear. Listeners provided a continuously changing rating of how well the probe fit the current context. Responses from Western listeners who were unfamiliar with scale structures reflected both an ability to extract statistical regularities in the Indian music as well as an attempt to assimilate the Indian music to tonal structures from the West (i.e., diatonic scales). Somewhat surprisingly, Western listeners adapted to tonal changes more rapidly than Indian listeners, perhaps because they used the statistical listening strategy described above.

Perception of Western pitch structures

Most cross-cultural research focuses on music that is 'foreign' to Western listeners. This makes practical sense: Music psychologists are generally based in Western nations, and Western music is widespread and thus generally well known. However, given historical claims about the importance of Western music, one must wonder whether individuals with no exposure to its pitch systems (e.g., equal-tempered tuning, diatonic scales) can learn, or even perhaps prefer, the kinds of structures used in the West.

It is a daunting task to answer such questions, because one must find a culture with no exposure to Western music. Recently, one research group led by Joshua McDermott did accomplish this task by traveling to a remote community in South America, the Tsimane', who can only be reached by canoe (McDermott, Schultz, Undurraga, & Godoy, 2016). In a series of experiments, this team tested both preference for consonance as well as the ability to discriminate consonant from dissonant chords. Importantly, the Tsimane' were just as adept at discriminating chords based on their consonance or dissonance as Westerners. So basic auditory perception functioned similarly across groups. However, the Tsimane' showed no preference for one over the other. This is similar to findings yielded by the aforementioned study of cotton-top tamarins (McDermott & Hauser, 2004), which also did not show a preference for consonance.

The dissociation among Tsimane' listeners between perceptual discrimination and preference is critical. First, it shows that a preference for consonance is probably learned rather than being a byproduct of basic perception, as was thought for many years (see chapter 3 for further discussion). Second, it suggests that the aforementioned result from tamarins may not reflect an impoverished perceptual system, but instead be based on the fact that their auditory environment does not use the consonant/dissonant distinction as a marker for pleasantness.

Can we follow the beat of another culture's rhythm?

Another fundamental way in which musical cultures differ is in rhythmic organization. Perhaps one of the most salient cross-cultural differences in rhythm comes from comparing drumming patterns from the Western classical tradition to African drumming. African drumming styles often do not suggest a simple metrical framework, which may confound the untutored listener. An example is shown in Figure 15.3. This pattern is common in West African drumming, and can be easily reproduced there, but is very difficult for most Western listeners (i.e., those with European ancestry) to learn and reproduce even though all the time intervals are related by a 2:1 serial ratio to each other (see chapter 6 for more discussion of serial ratios in rhythm).

What makes African rhythms sound so different from Western rhythms? One view from ethnomusicology is that African rhythms are organized differently from Western rhythms. Specifically, African rhythms may not be metrical in the same sense as Western music. Whereas Western listeners typically perceive rhythms as fitting into a metrical organization (see chapter 6), African rhythms are thought to be organized 'additively' (e.g., Stone, 1985). Metrical organization implies a hierarchy, in that smaller time-spans are expressed as ratios of larger time-spans. By contrast, ***additive timing*** (in this context) means that the organization

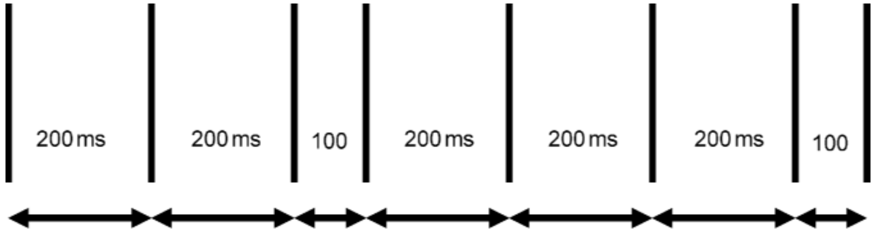


Figure 15.3 Time pattern common in West African drumming.

is serial rather than hierarchical. Each time-span in an additive organization is autonomous, not linked to a larger superordinate span. In practice, the additive organization of African drumming involves the use of a fixed rhythmic pattern that functions like a 'timeline,' which provides the basis for all other instruments (Jones, 1959). Often the timeline is played by a bell. Though the timeline functions similarly to a 'beat,' in that it establishes a referent time series that instrumentalists use, it is usually structured from unequal time-spans, like the pattern shown in Figure 15.3. In a Western sense, then, the 'beat' of African drumming is syncopated rather than steady. This state of affairs is perplexing to the Westerner who is used to equating the beat with regularity (see chapter 6).

Some empirical support for this view was reported by Magill and Pressing (1997). The authors recruited a master drummer who repeatedly produced a prototypical rhythmic pattern (shown in Figure 15.3). In a particularly creative twist, the authors devised competing mathematical models that could determine whether the drummer's pattern of timing variability matched the assumptions of additive timing from African musicology, or was more metrical, following Western musicology. Their starting point was the ***two-level timing model*** described in chapter 11 (Wing & Kristofferson, 1973). This model accounts for timing variability in simple rhythmic tasks (keeping the beat) as a confluence of two sources: an internal timekeeper and motor variability. This model is fundamentally non-metrical because it does not propose that performers conceptualize a special time-span for the beat (see chapter 6 for more on this concept).

Magill and Pressing adapted the two-level model for African drumming by allowing the internal clock to store a sequence of durations like those shown in Figure 15.3. The alternate model added an additional constraint that timing would be stabilized by a regular beat, which would reduce timing variability associated with that time-span. The authors found that it was not necessary to add such a constraint: The drummer's timing patterns were consistent with a version in which all timed events were produced independently, as in the 'timeline' analogy described above.

However, the idea that African meter is ametrical has been questioned by Temperley (2000). Based primarily on analyses of transcriptions made by Blacking (1973) and Jones (1959), Temperley argues that there is pervasive evidence that African drumming is guided by metrical principles that are consistent with Lerdahl and Jackendoff's *Generative Theory of Tonal Music* (1983; see chapters 6 and 7 in this volume). The striking differences between these musical traditions can be resolved by considering that syncopation plays a much larger role in African than in Western music. African listeners are in a sense more 'comfortable' with syncopation than are Western listeners. In other words, what appears to be a difference in fundamental organization (metric versus additive) may instead be a difference in terms of how much rhythms are allowed to deviate from a single universal organizational scheme (meter).

Cross-cultural differences in rhythm based on language

Beyond the salient differences between African and Western European rhythms, there are many differences across cultures with respect to conventional uses of rhythm. One potential source for these differences comes from rhythms used in language. Patel and Daniele (2003) reported a provocative correlation between the rhythmic structure of language and the rhythmic structure found in instrumental music. They focused on a distinction among languages based on rhythmic categories. In Europe, Germanic languages (which include English and Dutch) are often categorized as 'stress-timed,' whereas romance languages (such as French and Spanish) are 'syllable-timed' (Pike, 1945). ***Stress-timed*** languages make distinctions between strong and weak syllables that can be measured in syllable duration and that relate to how vowels are produced. For instance the 'a' in 'about' is produced differently than the 'a' in 'tall' because it is linguistically reduced when used in the unstressed first syllable of 'about.' As a result, patterns of syllable timing in a language like English alternate between long durations (stressed syllables) and short durations (unstressed syllables). Interestingly, instrumental music by English-speaking classical composers (like Edward Elgar, composer of *Pomp and Circumstance Marches*) similarly exhibits strong contrasts in the duration of adjacent tones.

In contrast to stress-timed languages, syllables in ***syllable-timed*** languages (such as French) are more equally timed. Consider the difference in the English pronunciation of 'decoration' with its uneven timing of syllables, and the more regular timing of the syllables in its French equivalent, *décoration*. French instrumental music (think Claude Debussy, composer of *Claire de Lune*) likewise exhibits greater regularity in the durations of adjacent tones than does English instrumental music. Patel and Daniele's (2003) findings point to the possibility that the rhythmic patterns used in one's native language influence the rhythms one generates while composing music.

Cross-cultural differences in metrical structure

Cross-cultural research has also addressed the closely related issue of meter. As discussed in chapter 6, meters can vary from simple to complex, with complex meters involving cycles that do not reduce to simple binary or ternary patterns. In Western music, these complex meters are rare, but they are common in other cultures. For instance, in South Indian classical music, the *mridangam* (a two-sided pitched hand drum, featured in Figure 15.4) interweaves patterns on a wide variety of metrical cycles called *talas* that range from the very short (3 beats per cycle) to the very long (116 beats per cycle). *Talas* do not have fixed accenting patterns, and this differentiates them from meters in Western music. Indian drumming features mathematical permutations and combinations in improvisatory embellishments as well as precomposed cadential figures that are conceived within various macro- and micro-subdivisions of the *tala* (Krishnamurthy, personal communication).

In chapters 6 and 8, we discussed research on the perception of complex meters that are common in Balkan music (Hannon & Trehub, 2005a). As noted in those chapters, the sense of meter can become deeply engrained in adulthood, although infant listeners appear to be equally able to learn rhythms that are presented within a simple and complex meter. From this perspective, it may be that children enter the world with universal predispositions for rhythm that over time develop into culturally specific tendencies. Subsequent research on adults further verified that the ability to perceive meter is more strongly based on the meters one is exposed to than the 'simplicity' of the meter (Hannon, Soley, & Ullal, 2012). In that study, listeners who were raised in Turkey were more adept at discriminating rhythms that were based on complex meters with a 7/8 time signature, which are prevalent in that culture, than Western listeners. However, Turkish listeners were poor at perceiving rhythms based on unfamiliar complex meters that had a 15/8 time signature.

In some cases, rhythmic organization may function as a better cross-cultural cue than pitch structure. For instance, research comparing European with Arabic listeners suggests that although both listeners were sensitive to rhythmic aspects of phrasing in Arabic instrumental improvisations, only Arabic listeners were sensitive to subtleties of the (modal) tonal system (Ayari & McAdams, 2003). Specifically, Arabic listeners perceived that a group boundary occurred when an instrumentalist switched from one mode to another within the

Figure 15.4 Acclaimed Indian percussionist Dr. Rohan Krishnamurthy plays the *mridangam*, the principal percussion instrument of South India with a history of over 2000 years.

Source: http://www.rohanrhythm.com/videos/; photograph used with permission of Rohan Krishnamurthy and RohanRhythm. Copyright © RohanRhythm.

Arabic modal system (*maqām*), as well as when the boundary was conveyed by changes in rhythmic grouping (see chapter 7). In contrast, Western listeners who were unfamiliar with Arabic modes were only sensitive to boundaries established by rhythmic groups.

How well do we remember music from another culture?

Music that is unfamiliar to us can sound unstructured, and different exemplars from an unfamiliar culture can be hard to distinguish. This phenomenon can even be observed in the small-scale cultures associated with different styles of music, like rock and classical. Ever heard someone complain that a given style of music 'just sounds like noise' or 'all sounds the same'?

These examples, if true on further inspection, have important implications for memory. Distinctiveness plays a critical role in memory (Hunt & McDaniel, 1993), and in order to distinguish items, it helps to relate them to an underlying structural schema. Thus, if music from another culture is truly heard as unstructured 'noise' (aside from one's preference, which that kind of statement usually conveys), then we should have a hard time recognizing music from that culture.

Several studies addressing this question have been carried out by Steven Demorest, Steven Morrison, and their colleagues. One study explored memory for music among listeners raised in the USA and in Turkey. Each group listened to music from their own culture, the other group's culture, or Chinese music, which was unfamiliar to both groups. The inclusion of Chinese music was important because Western music is typically familiar to listeners of other cultures, like Turkish listeners. In the task, listeners were first exposed to several musical excerpts from the different cultures. This was followed by a test phase in which listeners heard these excerpts again as well as excerpts from other musical pieces, and had to recognize which ones they heard previously. Listeners were more successful at recognizing music from their own culture than from another culture, a phenomenon referred to as an ***enculturation effect*** in memory (Demorest, Morrison, Beken, & Jungbluth, 2008). Importantly, the Turkish listeners were more accurate at recognizing Turkish music than music from the USA, despite having some familiarity with the latter. A subsequent neuroimaging study revealed that attempts to recognize culturally unfamiliar music were associated with greater activation in several brain areas associated with memory, reflecting the greater effort that such recognition requires (Demorest et al., 2010).

Do cultures perform music (and think about performance) differently?

Consider the prototypical solo performer of classical music: a well-dressed individual, alone on the stage, performing to a hushed crowed who sits in darkness until the end of the piece. One aspect of this scenario that is particularly important from a cultural perspective is the fact that the 'performer' is separated from the 'listener' at all! Western listeners tend to take this separation for granted. In other cultures, however, such an idea is considered rather strange. In most of the world, music-making has been a collective activity, with all persons participating in both the production and perception of music (cf. Schulkin, & Raglan, 2014). From this perspective, the fact that people are encouraged to stand and sing along at rock concerts should not be considered an oddity. What *is* odd is the practice of sitting quietly and silently at a classical music concert!

Another possible variation in the expression of music performance across cultures is that in music of many non-Western cultures, the integration of musical form with physical movement is closer than it is in Western music. For instance, John Baily (1985) analyzed the structure of Afghan music during a period in which a new stringed instrument, the Herati *dutãr* (a kind of lute), was developed. This 14-stringed instrument was conceptualized as a combination of the simple two-stringed *dutãr* and another 14-stringed instrument, the *rubãb*. In comparing the original two-stringed *dutãr* with the 14-stringed *rubãb*, Baily found that the structure of melodic movement and rhythmic organization for compositions from each instrument were constrained by the instrument design. For instance, music composed for the *rubãb* is characterized by a wider pitch range and more scalar movement (i.e., pitch motion in a continuing direction), presumably because the *rubãb* has more strings than the *dutãr*. Thus the correlation between range of motion a performer uses, and the range of pitch, may be limited by the physical boundaries that constrain motion.

Related claims about African music have been made, with the additional suggestion that the experience of music in this culture has as much to do with the movements that create sound as with the sound itself (Kubik, 1962). This can best be understood, perhaps, when one experiences a live performance of an African drumming ensemble, which invariably includes dance as well as music. In this context, dance is an integral part of the musical experience, not an addition to it. In chapter 9, we also saw how movement and music are closely interwoven in young children's early musical explorations and in many musical play forms such as their hand-clapping routines. Further, movement to music appears to be critical to the learning of rhythm and meter (chapter 8). The implication is that the Western 'art' music tradition of the quiet and immobile audience, refraining from spontaneous responses and applause until the end of the piece, may in fact be a counterproductive way to experience music.

As compelling as these analyses of non-Western music are, one might wonder how different these analyses are from *all* Western music. As stated above, the sharpest distinctions arise when comparing the music of Afghanistan and Africa with Western classical music. On the other hand, in Western popular and jazz music, it is common to associate movement with sound, thus suggesting greater commonality across these cultures. For instance, Iyer (2002) has claimed that timing patterns characteristic of jazz, such as the delay of a snare drum that follows the bass drum, are best understood when one conceptualizes jazz as a way of simulating movement. He cited as a specific example the practice of rhythmically alternating hand claps with stamping one's foot, a practice that can be found in a variety of African-based music. The pattern of drumbeats, in his view, embodies this manner of communicating rhythm with the body. Of course, one could argue that certain 'Western' genres of music such as rock and jazz have more direct musicological connections to Africa than to Europe, and as such do not truly constitute 'Western' forms.

Can we tell what emotion is conveyed by music of another culture?

A primary concern for persons interested in cross-cultural differences related to music is the question of emotional communication. As discussed in chapter 13, if music is said to have 'meaning,' part of that meaning is the communication of emotion. Author PQP recalls one of his first exposures to Chinese opera in Kaige Chen's 1993 film *Farewell My Concubine*. This film focuses on a specific musical theme (from an opera of the same name) that is repeated several times. It sounded very strange to the author during the first presentation – interesting, but not conveying any clear emotional message. By the end of the film, however, the music had become gradually more familiar. With that familiarity, the emotional contours of the theme became clearer and the tragic message of the theme came forth. It is plausible that such repeated exposure is necessary for the understanding of emotional communication from any foreign culture.

Or is it? In fact, research suggests that listeners can at least identify the emotion being communicated in the music of a different culture after little exposure, even though the listener may not be 'moved' internally by that music. In other words, this research focuses more on whether music *expresses* emotion, as opposed to whether it *induces* emotion (Eerola & Vuskoski, 2013, see chapter 14 for more on this distinction). For instance, Balkwill and Thompson (1999) played excerpts of Hindustani *ragas* to Western listeners (Canadian undergraduate students). Participants were asked to rate the *ragas* on the degree to which they expressed 'joy,' 'sadness,' 'anger,' and 'peace.' The Western students' ratings correlated strongly with the intended emotions of the ragas for three of the emotions: joy, sadness, and anger.

In another study, Balkwill, Thompson, and Matsunaga (2004) found that Japanese listeners accurately recognized the emotions that Japanese, Western, and Hindustani music were intended to convey – namely, joy, anger, or happiness. In both studies, Laura-Lee Balkwill and colleagues observed that listeners' judgments of emotions in music corresponded with listeners' evaluations of musical dimensions such as tempo, loudness, and complexity of the melody. For instance, music receiving high ratings for joy was usually judged to be fast in tempo and low in melodic complexity, whereas music perceived as conveying sadness was usually judged to be slow in tempo and high in melodic complexity.

In order to better explain these results, Balkwill and Thompson (1999) proposed the ***cue-redundancy model***, which is shown in Figure 15.5 (more detailed variations can be found in

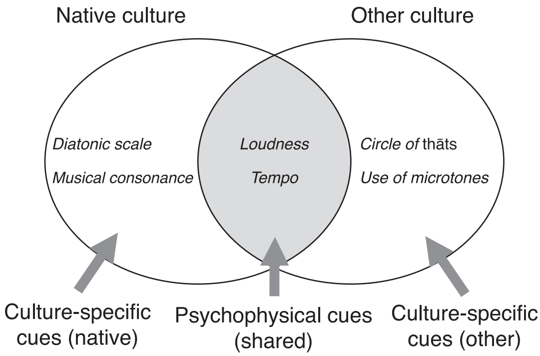


Figure 15.5 The cue-redundancy model.

Source: Adapted from Balkwill and Thompson (1999, Figure 1, p. 46) with permission from the authors and the University of California Press.

Fritz, 2012; Thompson & Balkwill, 2010). The basic point of models like the cue-redundancy model is that the meaning of certain musical features of music, thought of as 'cues' for a particular emotional interpretation, may not require culture-specific learning, whereas others do. An important implication of the previously discussed results is that the musical properties that communicate emotion are ***psychophysical cues*** that are shared across cultures, rather than those properties that are culture-specific. In some cases, these cues overlap, or are 'redundant.' For instance, tempo relates to a very basic physical quality: the rate at which information unfolds. The perception of tempo may not require immersion in a culture, given its association with a basic physical property. By contrast, tonal structures like the Western major or minor scale are based on subtle properties of pitch structure (see chapter 5), and thus may be based more on culturally specific knowledge.

A similar implication comes out of an earlier study by Meyer, Palmer, and Mazo (1998), who found that Western listeners identified the emotional intention of Russian laments as more sorrowful and more internally coherent if recordings included a specific timbral cue. In this case, the timbral cue was a 'gasping' sound that occurred when the lamenter breathed in (during an exaggerated and sustained bout of inhalation). Laments are characterized by long descending phrases in which the singer expresses his or her deep sorrow.

A limitation of the studies summarized so far is that participants often reported some prior exposure, however limited, to music that is intended to represent unfamiliar music. But what about responses to music in a tradition to which listeners have never been exposed? Fritz and colleagues (2009) studied a native African population (the Mafa) living in a very remote mountain region in the Cameroon, with no prior exposure to Western music at all. The researchers played excerpts of Western music and asked the Mafa participants to point to photographs of faces expressing different emotions to indicate what the music conveyed. The musical excerpts were specially composed for the experiment, and were designed to convey happiness, sadness, and fear. This was achieved by varying tempo, mode, pitch range, tone density, and rhythm in ways consistent with Western conventions for conveying these emotions, as found in previous research.

Fritz and colleagues (2009) found that the Mafa were able to identify three basic emotions (happiness, sadness, and fear) expressed by pieces of instrumental Western music. Mafa responses were driven in part by tempo, which is consistent with the cue-redundancy model, but somewhat surprisingly, they were also influenced by mode, which is a Western convention (though the authors noted that this pattern was 'more marked in Western listeners'; p. 573). With respect to mode, they classified most pieces in major key as happy, most minor pieces as scared/fearful, and pieces that were neither clearly major or minor as sad. Given the findings from developmental studies that mode is not one of the earliest cues used to judge emotion in music, it was surprising to find that listeners with no prior exposure to Western music seemed attentive to mode for judgments of emotion expressed in music. However, reliance on tempo and mode as cues for emotion expression was much more pronounced and consistent among Western listeners who served as a comparison group.

An important possibility to consider is that although musical properties vary greatly across cultures (e.g., tonal structure, rhythmic forms), there appears to be a universal set of emotional prototypes. This argument has been made strongly in a book entitled *The World in Six Songs* by Daniel Levitin (2008). According to Levitin, all cultures have produced types of songs that express important experiences held in common, which include *friendship*, *joy*, *comfort*, *knowledge*, *religion*, and *love*. By Levitin's account, music is an intrinsic part of human identity. Thus the messages communicated in music reflect core elements of humanity, and these elements come up as universal themes. Along similar lines, it has been observed by Sandra Trehub (2000) that music for children ubiquitously includes classes of lullabies and play songs, and that people can often recognize the distinction between these types of songs from another culture. Thus the close connection of music with one of the primary contexts of human life – the loving interactions between parent and child – begins right from the start.

Are there 'musical universals'?

It is common in summaries like this one to see authors offer a list of 'candidate' universal properties (e.g., Brown & Jordania, 2013, Carterette & Kendall, 1999; Justus & Hutsler, 2005). Many have been mentioned in passing here. The list often includes items like: the equivalence of pitches separated by the octave; organization of pitch into some kind of (typically asymmetric) tonal system; presence of the octave and perfect fifth in those tonal systems; the use of some kind of stable temporal referent (e.g., a pulse); and the segmentation of musical sequences into smaller groups or phrases. It is important to note the qualification here – nobody really knows which properties are universal, particularly given the unevenness with which different cultures are represented. Stevens and Byron (2016) make a point to focus on universal processes and leave music-specific properties out of consideration. Nonetheless, it is useful to consider whether a common core of properties may exist.

The general consensus seems to be that culturally universal properties have to do with general characteristics of musical structure, whereas culturally specific properties have to do with more detailed characteristics. For example, with respect to pitch, it has been proposed that a discrete pitch representation is universal, but the specific set of pitches varies. Similarly, asymmetric tonal scale systems may be universal, though the specific set of asymmetric intervals in a scale may vary, as we illustrated earlier. With respect to time, most seem to agree that synchronization to some kind of higher-order structure (like meter) is probably universal, as is some kind of alternating pattern of emphasis, but the specific structures to which one entrains and the form of alternation may vary, as in the West African 'timeline' discussed earlier.

Ultimately, of course, there is no easy 'yes' or 'no' answer to the question of universality. Music is intriguing in part because its meaning is simultaneously constrained by culture and yet extends across cultures. A comparison to language is informative. As with music, linguists typically believe that the world's languages comprise different specific manifestations of a common core of more general principles. However, music is unlike language in that the meaning of a foreign language is largely inaccessible to someone unfamiliar with that language. By contrast, research suggests that music's meaning can still be communicated to some degree to listeners when the music's culture is foreign to those listeners. At the same time, the intention behind music (as we discussed in chapter 13) is not fixed. Ultimately, it may be that music binds humanity together more effectively than does language, while at the same time acting as a vehicle for cultural diversity just like language does.

Final coda

As the title of this book implies, we are interested in how music as an auditory stimulus becomes a significant part of one's life, including the associations one has with a given culture. Our perspective is that the psychology of music encompasses the scope of musical experience from the physical vibrations of sound to the kind of deep significance that leads people to invest so much time, money, and emotional energy in music and music-making. Moreover, the ubiquity of music across cultures – regardless of its commonalities and differences – argues for the fundamental importance of music to human life.

Obviously, it is far too early to decide whether the significance of musical sound is a product of culturally learned norms or of universally shared neural and cognitive processes. It is more likely that music results from interactions between culturally specific and universal components. The most interesting course for research to follow is to explore the expression of universal ideas in different cultures.

Although it may be premature to draw generalizations from the existing research, some intriguing hypotheses for future research can be derived. First, it appears as though cultures share similar 'core' features, but embellish these core features in ways that please different listeners. Thus cultural specificity may best be found in varieties of musical complexity. Second, some research cited above suggests that more 'universal' characteristics of music can be found in the time domain, rather than in the pitch domain. Finally, with respect to performance, it appears that there are distinctions between 'art' music and 'folk' (or 'popular') music that exist across cultures. Moreover, cross-cultural comparisons need to be careful not to conflate 'culture' with musical style (e.g., classical versus popular). These ideas are, of course, tentative and suggestive, not conclusive.

And so we close with a speculation about music's overall significance. Along the lines of Levitin's (2008) proposal (and echoing earlier Socratic thinking), it might be said that music functions as a mirror of human experience. As noted in chapter 13, the way that music shapes our emotional response over time mirrors the intrinsically dynamic nature of how we experience emotion. Memories for significant events in our life are associated with music, such that the experience of a song (much like the experience of eating cakes soaked in tea described in Proust's *Remembrance of Things Past*) may transport us temporarily to the emotions of an earlier time in our life.

Moreover, to the degree that emotions are universal, certain elements of music may likewise be said to be universal. In this context, cultural differences in music may be analogous to differences across languages, with both serving to communicate universal ideas and feelings. Just as languages communicate common ideas with different words and syntax, so does music communicate similar ideas through different tonal and rhythmic structures. The profound significance of music and the immense joy that it has given people in all times and places is almost certainly the result of the power of music to reveal the intricacies and depths of human life.

Appendix: The chapters in action

Exercises and application assignments

Psychology of music is an inherently engaging topic, and lends itself to active learning. In this appendix, we provide some examples of brief exercises and application assignments to involve students in active learning, in tandem with chapter reading. We hope this will serve as a starting base for instructors to tailor their classes to fit, and to inspire further creative ideas to make the chapters come to life for students.

Brief exercises

The following brief exercises, suggested by Peter Pfordresher, are examples of brief demonstrations and straightforward conceptual replications of key findings that can be carried out with minimal resources.

Chapter 3 exercises

- Try the following demonstration on a friend, to explore localization of sound. Ask your friend to close his or her eyes. Stand about 1 yard (or 1 meter) away from your friend to one side and say his or her name. Ask your friend to point to where he or she thinks you are. Repeat this procedure several times, speaking at various positions in front of or behind your friend. Try high and low pitches, and bursts of sound such as snapping your fingers. Record how accurate your friend is at different locations, and consider how these results may reflect cues for localization discussed in chapter 3. Now ask your friend to repeat this procedure on you. Note that you may want to wear socks and walk on a carpeted floor, so the sound of your footsteps doesn't give away where you are moving!

- Ask a friend to stand as far away from you in a large room or hall as possible. Close your eyes and ask your friend to whisper one or two sentences. Try to repeat what was spoken. Then cup your outer ears (pinnae) with both hands, making sure your fingers are tightly pressed together and that there is no space between your fingers and thumb. Ask your friend to whisper again. Does extending your outer ears (pinnae) make a difference to your ability to decipher the speech? Does the whispering seem to sound closer to you? Reverse the roles.

Chapter 5 exercises

- Replicate the probe-tone technique described in chapter 5. On a piano, play the first seven notes of the C-major scale at a slow tempo. Then pause and play one of the 12 pitch chromas at random. Ask a friend to rate how well this final tone completed what you had performed on a scale from 1 to 7, where 7 indicates best possible completion. Repeat this procedure with all 12 chromas, and then compare your friend's ratings to the values below (from Krumhansl, 1990).

C	C#	D	D#	E	F	F#	G	G#	A	A#	B
6.35	2.23	3.48	2.33	4.38	4.09	2.52	5.19	2.39	3.66	2.29	2.88

- Record yourself singing the first line from one of your favorite songs. Then listen to a recording of that song. Were you within a semitone of the melody, as would be predicted by Levitin (1994), reviewed in chapter 5?

Chapter 6 exercises

- Set up your MP3 or CD player so that it's just about ready to play one of your favorite songs. Before you start the song, try to remember the tempo of the song's beat. Tap this beat on a surface. As you continue to tap, turn on the song. How close was the rate of your tapping to the song's actual tempo? Were you fairly close, as Levitin and Cook (1996) would predict?

- Listen to a song you know well in which the melody line is syncopated. (This can include almost any popular song – 'Let It Be' is a good example if you want to stay close to the book.) First, try tapping along to the beat that is established by the drum and/or musical instruments. Now try tapping along to the *accented syllable* of each sung word (e.g., 'MO-ther MA-ry COMES to ME'). Which is harder, tapping to the instruments or tapping to the words? Does your experience match the way in which lyrics are set to rock music, as found in the research by David Temperley (2000) discussed in chapter 6?

  Try it with some other pop songs. Do the accented syllables of the sung words tend to come *just prior to* the metrical pulse, thus drawing more attention to it? Does this seem to elicit more dance-like movement? Sing a few lines of the song out loud. How do your movements differ when singing the lyrics of the song more metrically versus in this syncopated fashion?

Chapter 11 exercises

- Ask a friend who is a musician to play a piece from his or her repertoire in three ways: in an expressively neutral or 'deadpan' fashion, in an exaggerated fashion, and as they would normally perform it. Listen to his or her performance and see how well you can distinguish these performances. What kind of 'cues' do you think you were using? Did you hear your friend use 'phrase final lengthening,' as described in chapter 11?

- Test your musical ability using one or more of the following online tests: the Seattle Singing Accuracy Protocol (https://ssap.music.northwestern.edu/), the Online Amusia Test (www.brams.org/en/onlinetest/), and the Profile of Music Perception Skills (https://www.uibk.ac.at/psychologie/fachbereiche/pdd/personality_assessment/proms/index.html.en). Based on your results, how would you categorize your ability? Do you think the tasks and manner of scoring on this test were a reasonable assessment of your ability? (This topic ties in with both chapters 9 and 11.)

Chapter 15 exercises

- Hindustani *ragas* are associated with different times of day and different emotions. There are websites with examples of *ragas* that describe these associations (e.g., http://raag-hindustani.com/Scales3.html). Play these for a friend who does not know Indian music, and test whether he or she can recognize any of these associations correctly. If so, ask your friend if he or she knows of any characteristics in the music that led to these conclusions.

- If you know anybody who has learned Indian classical music, repeat this procedure with that person and compare his or her answers with the friend who is not trained in this musical tradition. Ask your friend who knows Indian classical music to tell you more about how he or she engages with *ragas*.

Application assignments

In this section, we provide examples of assignments designed to engage students in active learning as they apply knowledge gleaned from each chapter. The parameters of each assignment are flexible, and the design can be tailored to classes of different sizes and academic levels, and to either individual or small group projects. Unless otherwise indicated, the following assignments were designed by Siu-Lan Tan.

Chapter 1 assignment (easily adapted to chapters 4–6, 8–12, 14, and 15)

'Scavenger Hunt': Peruse abstracts from the last 5–10 years in relevant journals in this field, as recommended by your professor, to find:

(a) five abstracts on topics that you have been interested to learn about before enrolling in this course;
(b) five abstracts on new topics that sound intriguing and that you have not encountered before;
(c) five abstracts outlining studies yielding results that are surprising to you (i.e., findings that you would not have predicted or intuited);
(d) three to five abstracts on one particular topic or general theme that shows up frequently within the last 5–10 years. There are many, so identify one thread that is of interest to you.

Finally, select one of the full journal articles that best fits your interests and level of background knowledge to read and outline. Print and read all the abstracts you collected in addition to the full article, and come to class ready to discuss what you found.

Flexible components: Instructors can tailor the guidelines for the search, the level of the selected journals, the number of abstracts to find, and the number of full articles the students are to read. Examples of journals suited for music education-oriented courses include *Journal of Research in Music Education*, *British Journal of Music Education*, and *Psychology of Music*. For more advanced students with a background in research methods, *Psychology of Music*, *Musicae Scientiae*, *Psychomusicology: Music, Mind and Brain*, and *Music Perception* are examples of fitting journals. This assignment can be adapted to many different chapters by narrowing the scope to a particular topic, for instance studies in music education (ch. 9), practice (ch. 10), topics in social psychology (ch. 12), emotion (ch. 14), and cross-cultural studies (ch. 15). For more advanced courses, the search may be aimed at neuroscientific studies (ch. 4), or studies pertaining to melody (ch. 5), rhythm (ch. 6), infant music cognition (ch. 8), or topics in performance (ch. 11).

Objective: To orient students to the current research and scope of the field (or of a particular area) in an engaging way. This is a good method for giving students an idea of the broader context before they hone in on particular studies.

Chapter 2 assignment

'Acoustic spaces of ______ (your college or university)': This assignment has been shared with a handful of colleagues, who have implemented it with their own variations, some of which we share below.

Your mission is to examine three acoustically interesting spaces of your choice on our campus (in groups of three to five members), drawing meaningful connections to chapter 2 after carefully studying it. Try to find three spaces that are very different from each other (with respect to size, shape, features, and construction materials), to avoid repetition. Be original and creative, with the aim of finding at least one space no other group has thought of! Examine the characteristics of each venue carefully, conduct your own informal sound tests, and draw as many specific connections to chapter 2 as you can. Bring at least two different instruments per group and your own copy of the book, to make specific connections to the text (recording page numbers, as you will need them for your report). Finally, submit a short report summarizing your observations with at least five to seven connections to the text for each venue (with page numbers). For each venue, include one photograph to capture the overall shape and dimensions of the space, and one photograph to capture a specific feature highlighted in your report.

Flexible components: Instructors can tailor this assignment to a group activity culminating in either individual or group reports, and vary the length and depth of the report. Matthew Bezdek at Stony Brook University provides a list of key terms from the book (e.g., direct/reflected sound, listener envelopment, horizontal and vertical clarity, etc.) to direct students' attention during the task. Michael Schutz at McMaster University asks groups to post their reports online so they can include audio recordings or short videos, and read each other's observations. Additional readings may expand on chapter 2 for courses emphasizing acoustics, such as articles in the special issue 'Performance Spaces for Music' of *Psychomusicology: Music, Mind, and Brain* (2015, vol. 25) on the occasion of the 100th birthday of Dr. Leo Beranek in 2014.

Objective: To observe how the various characteristics of a performance space shape our experience of sounds and music, by applying the concepts learned in chapter 2.

Chapter 8 assignment

'Science, simplified or sensationalized?': Find two or three blog posts or pages from non-academic websites in which you can spot at least three important inaccuracies or oversimplifications about the fetus' sensitivity to music and speech before birth. (Note that they need not be completely erroneous claims, but may simply be imprecise or overstated.) Print out the web pages, highlight the parts you will correct, and in a short paper, explain what points you think are imprecise or overstated and why. Use your knowledge from chapter 8 to provide more complete explanations (including page numbers from the book). Wherever you can, cite and outline a related study described in chapter 8. Finally, rephrase the sentences you highlighted to make them more accurate. Submit your short paper together with the printouts of the highlighted web pages.

Flexible components: Instructors can specify the how many web pages to include, how long critiques should be, and whether further independent reading (such as primary readings) beyond chapter 8 is required. This assignment can be adapted to any topic on which general misconceptions abound. Author ST also uses a variation of this structure for her exams, providing students with real web pages (on various topics in the book) to comment on and correct in a short essay question.

Objectives: To help students become more critical consumers of what they read on the web and in the popular media, and to deepen learning of chapter 8.

Chapter 8 assignment with media (can easily be adapted to most chapters)

'Music and the brain: Infant music cognition' (or other topics): Lecture 13 in neuroscientist Aniruddh Patel's (2015) *Music and the Brain* audio, or video lecture series in The Great Courses set, complements this chapter beautifully, providing an intriguing and clear presentation about music cognition in infants, focusing on singing and speech. Watch Lecture 13, entitled 'Development of music cognition' (30 minutes), and write a short report describing three important ways that the presentation expanded your understanding of the material presented in chapter 8 by providing a view from neuroscience (introduced in chapter 4 of this book). Alternatively, the series includes a companion guide and its own questions.

Flexible components: The other lectures in this set resonate with many other topics in this book, including presentations on consonance and dissonance (ch. 3), pitch and melody (ch. 5), rhythm (ch. 6), nature and nurture (ch. 10), music and language (ch. 4), music and emotion (ch. 14), and music and culture (ch. 15).

Objective: To expand students' knowledge about infant music cognition from the viewpoint of brain development, building on what they have learned in chapters 4 and 8.

Chapter 9 assignment (can also apply to chapter 8 content)

The following assignment was designed by Michael Schutz[1] at McMaster University.

'Sing, sing a song': Select four or five examples of infants and children singing in YouTube videos. Describe and discuss each video in detail, drawing as many connections to chapter 9 (and chapter 8, if you wish) as you can. Wherever possible, cite and outline a relevant study reviewed in the book to give more depth. Try to select videos that depict infants and children at different ages, to avoid repetition. In your closing paragraph, discuss links you see between the four or five videos. Together, what do they illustrate about the development of singing ability?

Flexible components: Instructors can tailor this assignment to fit an individual or group activity. Michael Schutz asks students to work in groups of four or five students, and posts the group reports along with the videos online, to foster communication between all the class members as they discuss each group's reports. For advanced courses, additional readings may be added to expand on chapter 9 by asking students to read some of the cited studies as primary sources. The authors of this book also recommend the article 'Mothers as singing mentors for infants' (Trehub & Gudmundsdottir, 2015) as engaging background reading that is accessible to newcomers, providing an outline of early singing development and eight links to videos discussed in the article.

Objective: To gain insight into the acquisition of singing in infants and young children.

Chapter 10 assignment

The following assignment was designed by Michael Schutz, who recommends it mainly for music students, and has implemented it with undergraduate students at McMaster University and graduate students at the University of Toronto.

'Best practices': Improve the effectiveness of your music practice by working in small groups (three or four students) to examine your approaches. Each group member should keep a daily practice log (typed) for seven consecutive days, recording the following:

Where: (location)
When: (how long in total, relative time for each segment)
What: (breakdown of material with approximate times, strategies used) *Students are encouraged to try approaches discussed in the textbook, especially if you have not often implemented them before.*
Why: (specific goals, areas of focus)

At the end of the week: (a) Each group member should produce a one-page (typed) reflection of the week's practice, including at least five specific connections to chapter 10 of the book (providing page numbers for each point). Wherever possible, make connections to a related study (describe the study and finding, include the page number where it appears, and explain how it is relevant to your practice). (b) In your group meeting, read and discuss each other's reports. Co-write and submit a one-page summary of what you learned from the discussion with your peers.

Objectives: To critically examine your current approaches to practicing, explore how to improve practice effectiveness and learn from your peers' experience.

Chapter 14 assignment

The following was co-designed by Matthew Bezdek (a former student of author ST who now teaches music cognition courses) and Siu-Lan Tan.

'The music of my life': This assignment will make you more aware of how music is integrated into your daily life. Unfortunately, for this class, we don't have the resources to hand out pagers such as the experience sampling method used by Sloboda and O'Neill (2001). Instead, keep a written log of every instance over the course of the next week when you hear any music. For each case, provide a general description of where you were, the source of the music, whether or not it was your choice to hear it, the title and artist if you know the piece (or use an app to try to identify it), the main characteristics of the music, the emotion or emotions expressed and/or induced by the music (if applicable), and any other items you wish to add. In your log, try to include audio or YouTube links to as many of the pieces of music as you can identify and find, as it will be a time capsule of the musical soundtrack for one week of your life. As an alternative to writing a log, you could try the free sampling MuPsych app (Randall and Rickard, 2017; see www.mupsych.com/research/).

Variation or extension: Record every time you hear music over the next week that you selected yourself. When and where did you listen to music? Why did you listen? Why did you select the particular piece(s)? Review the list of goals for music listening that Saarikallio (2011) gleaned from people ranging widely in age, or from other studies in chapter 14. Did your experience match some of these goals, or would you add some categories? Did your mood shift while listening to the music? Did your goals for emotion self-regulation change during your listening episodes?

Objectives: To help students become more aware of how much they are exposed to music, and gain insight into how and why they listen to music.

Chapter 14 assignment (can easily be adapted for chapters 8, 11, 12, and others)

'Research proposal: Film music' (or other topics): As apparent in the section on film music research in chapter 14, one standard way of studying the role of music in film is to present the same film scenes with different music tracks, in order to see how participants respond to the different pairings. After reading chapter 14 carefully, outline your own proposal for a research study using the 'switch-the-music-track' method, addressing your own original research question about film music. What is your hypothesis? Think critically about the design: Would you implement a between-subjects or within-subjects design, and why? Is it particularly important to present the film clips (and music) in different orders, and why? What particular film scenes and musical pieces would you employ, and why, and what are some inherent challenges to the choice of stimuli when using rich materials such as films and scores?

Flexible components: Instructors can tailor the assignment to classes with different levels of background in research methods, providing explanations where necessary, and adjusting the level of depth and detail expected in the proposal. If helpful as an accessible primer, author ST has written a chapter on the switch-the-music-track method for readers in the arts and humanities (see Tan, 2017a), and an introductory chapter on how music shapes the storyline of a scene (see Tan, 2017b). This assignment can easily be adapted to a range of topics after students have been familiarized with the standard methods in any area, such as the Goldberg paradigm (ch. 12), the habituation procedure (ch. 8), motion capture and point-light display (ch. 11), or any other methods outlined in the book.

Objectives: To help students formulate their own questions and design their own study on a topic of broad appeal, using a research paradigm they have learned.

Assignment for any chapters in the book

'What's the next question?': After reading the review of studies on a topic specified by your instructor in one of the book chapters, ask: What is the next question the researchers should ask? And how can it be studied (What method? Outline the procedure.). For Part A of your short report, type your responses to these questions. Then search the literature to see whether you can find a study that addresses the same or a similar question, how it was implemented, and learn what findings were yielded. If you cannot find one on your question, select a study that is at least relevant to your topic. In Part B of your short report, explain how you searched for and found the article (Which databases did you use? What key terms did you employ?). Finally, outline the method and main findings of the study, and submit a printout with your short paper.

Flexible components: This text is conducive to such an assignment because many topics are reviewed in some depth, methods are often explained, and many individual studies are outlined. This assignment can be adapted to many different topics reviewed in the book.

Objective: To train students to frame good research questions that emerge from a body of research, developing their problem-finding skills.

Note

1 The authors thank Michael Schutz for sharing assignments from his course. Guidelines are based closely on his handouts, and objectives are reproduced in his own words. The authors renamed the assignments in keeping with the style of the other assignments.

References

Abel, M. K. A., Li, H., Russo, F. A., Schlaug, G., & Loui, P. (2016). Audiovisual interval size estimation is associated with early musical training. *PLOS ONE*, *11*, e0163589.

Abeles, H. F. (2009). Are musical instrument gender associations changing? *Journal of Research in Music Education*, *57*, 127–139.

Abeles, H. F., Hoffer, C. R., & Klotman, R. H. (1994). *Foundations of music education* (2nd ed.). New York: Schirmer.

Abrams, R. M., Gerhardt, K., & Peters, A. J. M. (1995). Transmission of sound and vibration to the fetus. In J.-P. LeCanuet, W. Fifer, N. Krasnegor, & W. Smotherman (Eds.), *Fetal development: A psychobiological perspective* (pp. 315–330). Hillsdale, NJ: Lawrence Erlbaum.

Abrams, R. M., Griffiths, K., Huang, X., Sain, J., Langford, G., & Gerhardt, K. J. (1998). Fetal music perception: The role of sound transmission. *Music Perception*, *15*, 307–317.

Adachi, M., & Trehub, S. E. (1998). Children's expression of emotion in song. *Psychology of Music*, *26*, 133–153.

Adachi, M., & Trehub, S. E. (2012). Musical lives of infants. In G. E. McPherson & G. F. Welch (Eds.), *Oxford handbook of music education* (Vol. 1, pp. 229–247). Oxford: Oxford University Press.

Adachi, M., Trehub, S. E., & Abe, J.-I. (2004). Perceiving emotion in children's songs across age and culture. *Japanese Psychological Research*, *46*, 322–336.

Allen, R. (2013). Free improvisation and performance anxiety among piano students. *Psychology of Music*, *41*, 75–88.

Allport, G. (1954). *The nature of prejudice*. Cambridge, MA: Addison-Wesley.

Alluri, V., Toiviainen, P., Jääskeläinen, I. P., Glerean, E., Sams, M., & Brattico, E. (2012). Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm. *NeuroImage*, *59*, 3677–3689.

Altenmüller, E. (2005). Robert Schumann's focal dystonia. In J. Bogousslavsky & F. Boller (Eds.), *Neurological disorders in famous artists* (*Frontiers in neurology and neuroscience*, Vol. 19, pp. 1–10). Basel: Karger.

Altenmüller, E., Ioannou, C. I., & Lee, A. (2015). Apollo's curse: Neurological causes of motor impairments in musicians. *Progress in brain research*, *217*, 89–106.

American College of Obstetricians and Gynecologists (2013). *Definition of term pregnancy* (Committee Opinion No. 579). Retrieved from https://www.acog.org/-/media/Committee-Opinions/Committee-on-Obstetric-Practice/co579.pdf?dmc=1&ts=20161016T2020546367

American Heritage dictionary of the English language (2016). Acoustics, in *American Heritage dictionary of the English language* (5th updated hardcover ed.). New York: Houghton-Mifflin.

Ando, Y. (1985). *Concert hall acoustics*. Berlin: Springer.

Ashley, R. (2014). Expressiveness in funk. In D. Fabian, R. Timmers, & E. Schubert (Eds.), *Expressiveness in music performance: Empirical approaches across styles and cultures* (pp. 154–169). Oxford: Oxford University Press.

Ashley, R. (2016). *Musical improvisation*. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford handbook of music psychology* (2nd ed., pp. 667–679). Oxford: Oxford University Press.

Athanasopoulos, G., Tan, S.-L., & Moran, N. (2016). The influence of literacy on representation of time in musical stimuli: An exploratory cross-cultural study in Britain, Japan, and Papua New Guinea. *Psychology of Music*, *44*, 1126–1144.

Athos, E. A., Levinson, B., Kistler, A., Zemansky, J., Bostrom, A., Freimer, N., & Gitschler, J. (2007). Dichotomy and perceptual distortions in absolute pitch ability. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 14795–14800.

Axelsson, A., & Lindgren, F. (1981). Hearing in classical musicians. *Acta Otolaryngologica Supplement*, *377*, 3–74.

Ayari, M., & McAdams, S. (2003). Aural analysis of Arabic improvised instrumental music (Taqsīm). *Music Perception*, *21*, 159–216.

Ayotte, J., Peretz, I., & Hyde, K. (2002). Congenital amusia: A group study of adults afflicted with a music-specific disorder. *Brain*, *125*, 238–251.

Bachem, A. (1955). Absolute pitch. *Journal of the Acoustical Society of America*, *27*, 1180–1185.

Baddeley, A. D. (1986). *Working memory*. London: Oxford University Press.

Baharloo, S., Johnston, P. A., Service, S. K., Gitschier, J., & Freimer, N. B. (1998). Absolute pitch: An approach for identification of genetic and nongenetic components. *American Journal of Human Genetics*, *62*, 224–231.

Baharloo, S., Service, S. K., Risch, N. Gitschier, J., &. Freimer, N. B. (2000). Familial aggregation of absolute pitch. *American Journal of Human Genetics*, *67*, 755–758.

Baily, J. (1985). Music structure and human movement. In I. Cross, R. West, & P. Howell (Eds.), *Musical structure and cognition* (pp. 237–258). London: Academic Press.

Balkwill, L.-L., & Thompson, W. F. (1999). A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. *Music Perception*, *17*, 43–64.

Balkwill, L.-L., Thompson, W. F., & Matsunaga, R. (2004). Recognition of emotion in Japanese, Western, and Hindustani music by Japanese listeners. *Japanese Psychological Research*, *46*, 337–349.

Bangert, M., & Schlaug, G. (2006). Specialization of the specialized in features of external human brain morphology. *European Journal of Neuroscience*, *24*, 1832–1834.

Bangerter, A., & Heath, C. (2004). The Mozart effect: Tracking the evolution of scientific legend. *British Journal of Psychology*, *43*, 605–623.

Baumeister, R. E. (1984). Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology*, *46*, 610–620.

Beament, J. (1997). *The violin explained*. Oxford: Oxford University Press.

The Beatles (1967). *Sgt. Pepper's Lonely Hearts Club Band*. London: EMI Records.

Beilock, S. L. (2010). *Choke: What the secrets of the brain reveal about getting it right when you have to*. New York: Simon & Schuster.

Beineix, J.-J. (Director), & Ossard, C. (Producer). (1981). *Diva* [Motion picture]. Paris: Les Films Galaxie. (Available from Anchor Bay Entertainment, 1699 Stutz Drive, Troy, MI 48094)

Bengtsson, S. L., Ullén, F., Henrik Ehrsson, H., Hashimoto, T., Kito, T., Naito, E., … Sadato, N. (2009). Listening to rhythms activates motor and premotor cortices. *Cortex*, *45*, 62–71.

Benson, D. (2006). *Music: A mathematical offering*. Cambridge: Cambridge University Press.

Beranek, L. L. (2004). *Concert halls and opera houses: Music, acoustics and architecture*. New York: Springer.

Beranek, L. L. (2007, October). *Aspects of concert hall acoustics*. Richard C. Heyser Memorial Lecture presented at the *123rd Convention of the Audio Engineering Society*. Retrieved from www.aes.org/technical/heyser/downloads/AES123heyser-Beranek.pdf

Beranek, L. L. (2015). Concert hall design: Some considerations. *Psychomusicology: Music, Mind, & Brain*, *25*, 181–186.

Berger, A. A., & Cooper, S. (2003). Musical play: A case study of preschool children and parents. *Journal for Research in Music Education*, *51*, 151–165.

Bergeson, T. R. (2012). Spoken language development in infants who are deaf or hard of hearing: The role of maternal infant-directed speech. *The Volta Review*, *112*, 171–180.

Bergeson, T. R., & Trehub, S. E. (2002). Absolute pitch and tempo in mothers' songs to infants. *Psychological Science*, *13*, 72–75.

Berliner, P. F. (1994). *Thinking in jazz: The infinite art of improvisation*. Chicago, IL: University of Chicago Press.

Bernardi, N. F., Schories, A., Jabusch, H.-C., Colombo, B., & Altenmüller, E. (2013). Mental practice in music memorization: An ecological-empirical study. *Music Perception*, *30*, 275–290.

Bernstein, L. (1976). *The unanswered question: Six talks at Harvard*. Cambridge, MA: Harvard University Press.

Bever, T., & Chiarello, R. J. (1974, August 9). Cerebral dominance in musicians and nonmusicians. *Science*, *185*, 537–539.

Bharucha, J. J. (1984). Anchoring effects in music: The resolution of dissonance. *Cognitive Psychology*, *16*, 485–518.

Bhatara, A., Tirovolas, A. K., Duan, L. M., Levy, B., & Levitin, D. J. (2011). Perception of emotional expression in musical performance. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 921–934.

Bigand, E., Vieillard, S., Madurell, F., Marozeau, J., & Dacquet, A. (2005). Multidimensional scaling of emotional responses to music: The effect of musical expertise and duration of the excerpts. *Cognition & Emotion*, *19*, 113–139.

Bishop, L., Bailes, F., & Dean, R. T. (2013). Musical imagery and the planning of dynamics and articulation during performance. *Music Perception*, *31*, 97–117.

Bissonnette, J., Dubé, F., Provencher, M. D., & Moreno Sala, M. T. (2015). Virtual reality exposure training for musicians: Its effects on performance anxiety and quality. *Medical Problems of Performing Artists*, *30*, 169–177.

Bjørkvold, J. (1992). *The music within: Creativity and communication, song and play from childhood through maturity* (W. H. Halverson, Trans.). New York: HarperCollins.

Black, J. W. (1951). The effect of delayed side-tone upon vocal rate and intensity. *Journal of Speech and Hearing Disorders*, *16*, 56–60.

Blacking, J. (1973). *How musical is man?* Seattle, WA: University of Washington.

Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences*, *98*, 11818–11823.

Blood, A. J., Zatorre, R. J., Bermudez, P., & Evans, A. C. (1999). Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nature Neuroscience*, *2*, 382–387.

Boerner, S., & von Streit, C. F. (2007). Promoting orchestral performance: The interplay between musicians' mood and a conductor's leadership style. *Psychology of Music*, *35*, 132–143.

Bolivar, V. J., Cohen, A. J., & Fentress, J. C. (1994). Semantic and formal congruency in music and motion pictures: Effects on the interpretation of visual action. *Psychomusicology*, *13*, 28–59.

Bolton, T. L. (1894). Rhythm. *American Journal of Psychology*, *6*, 145–238.

Boltz, M. G. (2001). Musical soundtracks as a schematic influence on the cognitive processing of filmed events. *Music Perception*, *18*, 427–454.

Boltz, M. G. (2004). The cognitive processing of film and musical soundtracks. *Memory & Cognition*, *32*, 1194–1205.

Boltz, M. G. (2013). Music videos and visual influences on music perception and appreciation: Should you want your MTV? In S.-L. Tan, A. J. Cohen, R. A. Kendall, & S. D. Lipscomb (Eds.), *Psychology of music in multimedia* (pp. 217–234). Oxford: Oxford University Press.

Boltz, M., & Jones, M. R. (1986). Does rule recursion make melodies easier to reproduce? If not, what does? *Cognitive Psychology*, *18*, 389–431.

Boltz, M., Schulkind, M., & Kantra, S. (1991). Effects of background music on the remembering of filmed events. *Memory & Cognition*, *19*, 595–606.

Bonneville-Roussy, A., & Bouffard, T. (2015). When quantity is not enough: Disentangling the roles of practice time, self-regulation and deliberate practice in musical achievement. *Psychology of Music*, *43*, 686–704.

Boone, R. T., & Cunningham, J. G. (1998). Children's decoding of emotion in expressive body movement: The development of cue attunement. *Developmental Psychology*, *34*, 1007–1016.
Bowling, D. L. (2013). A vocal basis for the affective character of musical mode in melody. *Frontiers in Psychology*, *4*, 464.
Bowling, D., Sundararajan, J., Han, S., & Purves, D. (2012). Expression of emotion in Eastern and Western music mirrors vocalization. *PLOS ONE*, *7*, e31942.
Brackbill, Y., Adams, G., Crowell, D. H., & Gray, M. L. (1966). Arousal level in neonates and preschool children under continuous auditory stimulation. *Journal of Experimental Child Psychology*, *4*, 178–188.
Bradshaw, E., & McHenry, M. A. (2005). Pitch discrimination and pitch matching abilities of adults who sing inaccurately. *Journal of Voice*, *19*, 431–439.
Bramley, S., Dibben, N., & Rowe, R. (2016). Investigating the influence of music tempo on arousal and behaviour in laboratory virtual roulette. *Psychology of Music*, *44*, 1389–1403.
Brattico, E., Tervaniemi, M., Näätänen, R., & Peretz, I. (2006). Musical scale properties are automatically processed in the human auditory cortex. *Brain Research*, *1117*, 162–174.
Bregman, A. S. (1990). *Auditory scene analysis*. Cambridge, MA: MIT Press.
Bregman, M. R., Patel, A. D., & Gentner, T. Q. (2016). Songbirds use spectral shape, not pitch, for sound pattern recognition. *Proceedings of the National Academy of Sciences (USA)*, *113*, 1666–1671.
Breinbauer, H. A., Anabalón, J. L., Gutierrez, D., Cárcamo, R., Olivares, C., & Caro, J. (2012). Output capabilities of personal music players and assessment of preferred listening levels of test subjects: Outlining recommendations for preventing music-induced hearing loss. *Laryngoscope*, *122*, 2549–2556.
Brochard, R., Abecasis, D., Potter, D., Ragot, R., & Drake, C. (2003). The 'ticktock' of our internal clock: Direct brain evidence of subjective accents in isochronous sequences. *Psychological Science*, *14*, 362–366.
Broesch, T. L., & Bryant, G. A. (2015). Prosody in infant-directed speech is similar across Western and traditional cultures. *Journal of Cognition and Development*, *16*, 31–43.
Brophy, T. S. (2005). A longitudinal study of selected characteristics of children's melodic improvisations. *Journal of Research in Music Education*, *53*, 120–133.
Brown, J. C., Houix, O., & McAdams, S. (2001). Feature dependence in the automatic identification of musical woodwind instruments. *Journal of the Acoustical Society of America*, *109*, 1064–1072.
Brown, R. M., & Palmer, C. (2012). Auditory-motor learning influences auditory memory for music. *Memory & Cognition*, *40*, 567–578.
Brown, R. M., Zatorre, R. J., & Penhune, V. B. (2015). Expert music performance: Cognitive, neural, and developmental bases. *Progress in Brain Research*, *217*, 57–86.
Brown, S. (2000). The 'musilanguage' model of music evolution. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 271–300). Cambridge, MA: MIT Press.
Brown, S. (2006). The perpetual music track: The phenomenon of constant musical imagery. *Journal of Consciousness Studies*, *13*, 25–44.
Brown, S., & Jordania, J. (2013). Universals in the world's musics. *Psychology of Music*, *41*, 229–248.
Brown, S., Ngan, E., & Liotti, M. (2008). A larynx area in the human motor cortex. *Cerebral Cortex*, *18*, 837–845.
Bullerjahn, C., & Güldenring, M. (1994). An empirical investigation of effects of film music using qualitative content analysis. *Psychomusicology*, *13*, 99–118.
Burnham, D., Kitamura, C., & Vollmer-Conna, U. (2002). What's new, pussycat? On talking to babies and animals. *Science*, *296*, 1435.
Burns, E. M., & Ward, W. D. (1978). Categorical perception – phenomenon or epiphenomenon: Evidence from experiments in the perception of melodic musical intervals. *Journal of the Acoustical Society of America*, *63*, 456–458.
Butler, C. (2004). *Pleasure and the arts*. Oxford: Oxford University Press.

Cameron, J. E., Duffy, M., & Glenwright, B. (2015). Singers take center stage! Personality traits and stereotypes of popular musicians. *Psychology of Music*, *43*, 818–830.

Cariani, P. A., & Delgutte, B. (1996). Neural correlates of the pitch of complex tones: Pitch and pitch salience. *Journal of Neurophysiology*, *76*, 1698–1716.

Carter, C. E., & Grahn, J. A. (2016). Optimizing music learning: Exploring how blocked and interleaved practice schedules affect advanced performance. *Frontiers in Psychology*, 7, 1251.

Carterette, E. C., & Kendall, R. A. (1999). Comparative music perception and cognition. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 725–792). San Diego, CA: Academic Press.

Cassidy, J. W., & Standley, J. M. (1995). The effect of music listening on physiological responses of premature infants in the NICU. *Journal of Music Therapy, Special Children's Issue*, *32*, 208–227.

Castellano, G., Mortillaro, M., Camurri, A., Volpe, G., & Scherer, K. (2008). Automated analysis of body movement in emotionally expressive piano performances. *Music Perception*, *26*, 103–120.

Castellano, M. A., Bharucha, J. J., & Krumhansl, C. L. (1984). Tonal hierarchies in the music of North India. *Journal of Experimental Psychology: General*, *113*, 394–412.

Castro, S. L., & Lima, C. F. (2014). Age and musical expertise influence emotion recognition in music. *Music Perception*, *32*, 125–142.

Céline, M., & Trainor, L. J. (2013). Development of simultaneous pitch encoding: infants show a high voice superiority effect. *Cerebral Cortex*, *23*, 660–669.

Cevasco, A. M. (2008). The effects of mothers' singing on full-term and preterm infants and maternal emotional responses. *Journal of Music Therapy*, *45*, 273–306.

Chabris, C. F. (1999). Prelude or requiem for the 'Mozart effect'? *Nature*, *400*, 826–827.

Chaffin, R. (2007). Learning *Clair de Lune*: Retrieval practice and expert memorization. *Music Perception*, *24*, 377–393.

Chaffin, R., Demos, A. P., & Logan, T. R. (2016). Performing from memory. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford handbook of music psychology* (2nd ed., pp. 559–572). Oxford: Oxford University Press.

Chaffin, R., Imreh, G., & Crawford, M. (2002). *Practicing perfection: Memory and piano performance*. Mahwah, NJ: Lawrence Erlbaum.

Chaffin, R., Lisboa, T., Logan, T., & Begosh, K. T. (2010). Preparing for memorized cello performance: The role of performance cues. *Psychology of Music*, *38*, 3–30.

Chan, C., & Ackermann, B. (2014). Evidence-informed physical therapy management of performance-related musculoskeletal disorders in musicians. *Frontiers in Psychology*, *5*, 706.

Chan, L. P., Livingstone, S. R., & Russo, F. A. (2013). Facial mimicry in response to song. *Music Perception*, *30*, 361–367.

Chang, H-.W., & Trehub, S. E. (1977). Auditory processing of relational information by young infants. *Journal of Experimental Child Psychology*, *24*, 324–331.

Chapin, H., Jantzen, K., Kelso, J. A. S., Steinberg, F., & Large, E. (2010). Dynamic emotional and neural responses to music depend on performance expression and listener experience. *PLOS ONE*, *5*, e13812.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*, 893–910.

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex*, *18*, 2844–2854.

Chen, J., Woollacott, M. H., Pologe, S., & Moore, G. P. (2008). Pitch and space maps of skilled cellists: Accuracy, variability, and error correction. *Experimental Brain Research*, *188*, 493–503.

Chen, L., Zhou, S., & Bryant, J. (2007). Temporal changes in mood repair through music consumption: Effects of mood, mood salience, and individual differences. *Media Psychology*, *9*, 695–713.

Chion, M. (1994). *Audio-vision: Sound on screen* (C. Gorbman, Ed. & Trans.). New York: Columbia University Press.

Chomsky, N. (1966). *Cartesian linguistics*. New York: Harper and Row.

Cirelli, L. K., Einarson, K. M., & Trainor, L. J. (2014). Interpersonal synchrony increases prosocial behavior in infants. *Developmental Science*, *17*, 1003–1011.

Cirelli, L. K., Wan, S. J., & Trainor, L. J. (2016). Social effects of movement synchrony: Increased infant helpfulness only transfers to affiliates of synchronously moving partners. *Infancy*, *21*, 807–821.

Clarke, E. F. (1987). Categorical rhythm perception: An ecological perspective. In A. Gabrielsson (Ed.), *Action and perception in rhythm and music* (pp. 19–33). Stockholm: Royal Swedish Academy of Music.

Clarke, E., & Baker-Short, C. (1987). The imitation of perceived rubato: A preliminary study. *Psychology of Music*, *15*, 58–75.

Clifton, R. K., Perris, E. E., & Bullinger, A. (1991). Infants' perception of auditory space. *Developmental Psychology*, *27*, 187–197.

Coffman, D. D. (1990). Effects of mental practice, physical practice, and knowledge of results on piano performance. *Journal of Research in Music Education*, *38*, 187–196.

Cohen, A. J. (2009). Music in performance arts: Film, theatre and dance. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford handbook of music psychology* (pp. 441–451). Oxford: Oxford University Press.

Cohen, A. J. (2013). Congruence-Association Model of music and multimedia: Origin and evolution. In S.-L. Tan, A. J. Cohen, S. D. Lipscomb, & R. A. Kendall (Eds.), *The psychology of music in multimedia* (pp. 17–47). Oxford: Oxford University Press.

Colley, A. M., North, A. C., & Hargreaves, D. J. (2003). Gender bias in the evaluation of new age music. *Scandinavian Journal of Psychology*, *44*, 125–131.

Conklin, N. M. (2011). *Musical performance anxiety in virtual performances: A comparison of recorded and live performance contexts* (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses Open (3453579).

Cook, N. (1987). The perception of large-scale tonal closure. *Music Perception*, *5*, 197–206.

Cook, P., Rouse, A., Wilson, M., & Reichmuth, C. J. (2013). A California sea lion (*Zalophus californianus*) can keep the beat: Motor entrainment to rhythmic auditory stimuli in a non vocal mimic. *Journal of Comparative Psychology*, *127*, 1–16.

Cooke, D. (1959). *The language of music*. Oxford: Oxford University Press.

Corrigall, K. A., & Schellenberg, E. G. (2015). Predicting who takes music lessons: Parent and child characteristics. *Frontiers in Psychology*, *6*, 282.

Corrigall, K. A., Schellenberg, E. G., & Misura, N. M. (2013). Music training, cognition, and personality. *Frontiers in Psychology*, *4*, 222.

Corrigall, K., & Trainor, L. J. (2009). Effects of musical training on key and harmony perception. *Annals of the New York Academy of Sciences*, *1169*, 164–168.

Costa-Giomi, E. (2004). Effects of three years of piano instruction on children's academic achievement, school performance, and self-esteem. *Psychology of Music*, *32*, 139–152.

Cottrell, N. B. (1972). Social facilitation. In C. G. McClintock (Ed.), *Experimental social psychology* (pp. 185–235). New York: Holt, Rinehart & Winston.

Cottrell, N. B., Wack, D. L., Sekerak, G. J, & Rittle, R. H. (1968). Social facilitation of dominant responses by the presence of an audience and the mere presence of others. *Journal of Personality and Social Psychology*, *9*, 245–250.

Cowan, N. (2015). George Miller's magical number of immediate memory in retrospect: Observations on the faltering progression of science. *Psychological Review*, *122*, 536–541.

Črnčec, R., Wilson, S. J., & Prior, M. (2006). No evidence for the Mozart effect in children. *Music Perception*, *23*, 305–317.

Cross, I. (2003). Music, cognition, culture, and evolution. In I. Peretz & R. Zatorre (Eds.), *The cognitive neuroscience of music* (pp. 42–56). New York: Oxford University Press.

Cross, I. (2005). Music and meaning, ambiguity, and evolution. In D. Miell, R. MacDonald, & D. J. Hargreaves (Eds.), *Musical communication* (pp. 27–43). Oxford: Oxford University Press.

Cross, I., & Tolbert, E. (2016). Music and meaning. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford handbook of music psychology* (2nd ed., pp. 33–46). Oxford: Oxford University Press.

Cuddy, L. L., & Badertscher, B. (1987). Recovery of the tonal hierarchy – some comparisons across age and levels of musical experience. *Perception and Psychophysics*, *41*, 609–620.

Cuddy, L. L., Balkwill, L., Peretz, I., & Holden, R. R. (2005). Musical difficulties are rare: A study of 'tone deafness' among university students. *Annals of the New York Academy of Sciences*, *1060*, 311–324.

Cuddy, L. L., Cohen, A. J., & Mewhort, D. J. K. (1981). Perception of structure in short melodic sequences. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 869–883.

Cuddy, L. L., & Lunney, C. A. (1995). Expectancies generated by melodic intervals: Perceptual judgments of melodic continuity. *Perception & Psychophysics*, *57*, 451–462.

Curtis, M. E., & Bharucha, J. J. (2010). The minor third communicates sadness in speech, mirroring its use in music. *Emotion*, *10*, 335–348.

Custodero, L. A. (2002). The musical lives of young children: Inviting, seeking, and initiating. *Zero-to-Three Bulletin*, *23*, 4–9.

Custodero, L. A. (2006). Singing practices in 10 families with young children. *Journal of Research in Music Education*, *54*, 37–56.

Dahl, S., & Friberg, A. (2007). Visual perception of expressiveness in musicians' body movements. *Music Perception*, *24*, 433–455.

Dalla Bella, S. (2016). Music and brain plasticity. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford handbook of music psychology* (2nd ed., pp. 325–342). Oxford: Oxford University Press.

Dalla Bella, S., Giguère, J.-F., & Peretz, I. (2007). Singing proficiency in the general population. *Journal of the Acoustical Society of America*, *121*, 1182–1189.

Dalla Bella, S., Giguère, J.-F., & Peretz, I. (2009). Singing in congenital amusia. *Journal of the Acoustical Society of America*, *126*, 414–424.

Dalla Bella, S., Peretz, I., & Aronoff, N. (2003). Time course of melody recognition: A gating paradigm study. *Perception & Psychophysics*, *65*, 1019–1028.

Dalla Bella, S., Peretz, I., Rousseau, L., & Gosselin, N. (2001). A developmental study of the affective value of tempo and mode in music. *Cognition*, *80*, B1–B10.

Davidson, J. W. (1993). Visual perception of performance manner in the movements of solo musicians. *Psychology of Music*, *21*, 103–113.

Davidson, J. W. (2005). Bodily communication in music performance. In D. Miell, R. MacDonald, & D. J. Hargreaves (Eds.), *Musical communication* (pp. 215–237). Oxford: Oxford University Press.

Davidson, J. W. (2012). Bodily movement and facial actions in expressive musical performance by solo and duo instrumentalists: Two distinctive case studies. *Psychology of Music*, *40*, 595–633.

Davidson, L. (1994). Songsinging by young and old: A developmental approach to music. In R. Aiello & J. A. Sloboda (Eds.), *Musical perceptions* (pp. 99–130). New York: Oxford University Press.

Davidson, L., McKernon, P., & Gardner, H. (1981). The acquisition of song: A developmental approach. In J. A. Mason (Ed.), *Documentary report of the Ann Arbor Symposium on the Application of Psychology to the Teaching and Learning of Music* (pp. 301–315). Reston, VA: Music Education National Conference.

Davies, C. (1994). The listening teacher: An approach to the collection and study of invented songs of children aged 5 to 7. In H. Lees (Ed.), *Musical connections: Tradition and change* (pp. 120–127). Auckland: International Society for Music Education.

Davies, J. B. (1978). *The psychology of music*. Stanford, CA: Stanford University Press.

Davis, S. (2006). Implied polyphony in the solo string works of J. S. Bach: A case for the perceptual relevance of structural expression. *Music Perception*, *23*, 423–446.

Deaville, J. A., Rodman, R., & Tan, S.-L. (In progress). *The Oxford handbook of music and advertising*. New York: Oxford University Press.

DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science*, *208*, 1174–1176.

DeCasper, A. J., & Spence, M. J. (1986). Prenatal maternal speech influences new-borns' perception of speech sounds. *Infant Behavior & Development*, *9*, 133–150.

de l'Etoile, S. K. (2006). Infant-directed singing: A theory for clinical intervention. *Music Therapy Perspectives*, *24*, 22–29.

de l'Etoile, S. K., & Leider, C. N. (2011). Acoustic parameters of infant-directed singing in mothers with depressive symptoms. *Infant Behavior and Development*, *34*, 248–256.

Deliège, I. (1987). Grouping conditions in listening to music: An approach to Lerdahl & Jackendoff's grouping preference rules. *Music Perception*, *4*, 325–360.

Demorest, S. M., Morrison, S. J., Beken, M. N., & Jungbluth, D. (2008). Lost in translation: An enculturation effect in music memory performance. *Music Perception*, *25*, 213–223.

Demorest, S. M., Morrison, S. J., Stambaugh, L. A., Beken, M., Richards, T., & Johnson, C. (2010). An fMRI investigation of the cultural specificity of music memory. *Social, Cognitive & Affective Neuroscience*, *5*, 282–291.

Demorest, S. M., & Pfordresher, P. Q. (2015). Singing accuracy development from K–adult: A comparative study. *Music Perception*, *32*, 293–302.

Desain, P., & Honing, H. (1994). Does expressive timing in music performance scale proportionally with tempo? *Psychological Research*, *56*, 285–292.

Desain, P., & Honing, H. (2003). The formation of rhythmic categories and metric priming. *Perception*, *32*, 341–365.

Deutsch, D. (1972, March 3). Mapping of interactions in the pitch memory store. *Science*, *175*, 1020–1022.

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

Deutsch, D. (1995). *Musical illusions and paradoxes* [CD]. La Jolla, CA: Philomel Records.

Deutsch, D. (2013). Grouping mechanisms in music. In D. Deutsch (Ed.), *The psychology of music* (3rd ed., pp. 183–248). San Diego, CA: Academic Press.

Deutsch, D., Dooley, K., Henthorn, T., & Head, B. (2009). Absolute pitch among students in an American music conservatory: Association with tone language fluency. *Journal of the Acoustical Society of America*, *125*, 2398–2403.

Deutsch, D., Henthorn, T., Marvin, E., & Xu, H. (2006). Absolute pitch among American and Chinese conservatory students: Prevalence differences, and evidence for a speech-related critical period. *Journal of the Acoustical Society of America*, *119*, 719–722.

DeWitt, L. A., & Samuel, A. G. (1990). The role of knowledge-based expectations in music perception: Evidence from musical restoration. *Journal of Experimental Psychology: General*, *119*, 123–144.

Dickey, M. R. (1991). A comparison of verbal instruction and nonverbal teacher–student modeling in instrumental ensembles. *Journal of Research in Music Education*, *39*, 132–142.

Dissanayake, E. (2000). Antecedents of the temporal arts in early mother–infant interaction. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 389–410). Cambridge, MA: MIT Press.

Doane, M. A. (1980). The voice in the cinema: The articulation of body and space. *Yale French Studies*, *60*, 34.

Doheny, L., Hurwitz, S., Insoft, R., Ringer, S., & Lahav, A. (2012). Exposure to biological maternal sounds improves cardiorespiratory regulation in extremely preterm infants. *Journal of Maternal-Fetal & Neonatal Medicine*, *25*, 1591–1594.

Dowling, W. J. (1968). *Rhythmic fission and the perceptual organization of tone sequences*. Unpublished doctoral dissertation, Harvard University.

Dowling, W. J. (1973). The perception of interleaved melodies. *Cognitive Psychology*, *5*, 322–337.

Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, *85*, 341–354.

Dowling, W. J. (1984). Development of musical schemata in children's spontaneous singing. In W. R. Crozier & A. J. Chapman (Eds.), *Cognitive processes in the perception of art* (pp. 145–164). Amsterdam: North-Holland.

Dowling, W. J., & Bartlett, J. C. (1981). The importance of interval information in long-term memory for melodies. *Psychomusicology*, *1*, 30–49.

Dowling, W. J., Kwak, S., & Andrews, M. W. (1995). The time course of recognition of novel melodies. *Perception & Psychophysics*, *57*, 136–149.

Draganova, R., Eswaran, H., Murphy, P., Huotilainen, M., Lowery, C., & Preissl, H. (2005). Sound frequency change detection in fetuses and newborns, a magnetoencephalographic study. *NeuroImage*, *28*, 354–361.

Drake, C., & Botte, M.-C. (1993). Tempo sensitivity in auditory sequences: Evidence for a multiple-look model. *Perception & Psychophysics*, *54*, 277–286.

Drake, C., Jones, M. R., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: Attunement, referent period, focal attending. *Cognition*, *77*, 251–288.

Drake, C., & Palmer, C. (2000). Skill acquisition in music performance: Relations between planning and temporal control. *Cognition*, *74*, 1–32.

Drake, C., Penel, A., & Bigand, E. (2000). Tapping in time with mechanically and expressively performed music. *Music Perception*, *18*, 1–25.

Drennan, W. R., & Rubinstein, J. T. (2008). Music perception in cochlear implant users and its relationship with psychophysical capabilities. *Journal of Rehabilitation Research & Development*, *45*, 779–789.

Duke, R. A., Simmons, A. L., & Cash, C. D. (2009). It's not how much; it's how: Characteristics of practice behavior and retention of performance skills. *Journal of Research in Music Education*, *56*, 310–321.

Edwards, J. (2011). *Music therapy and parent–infant bonding.* Oxford: Oxford University Press.

Eerola, T., Luck, G., & Toiviainen, P. (2006). An investigation of pre-schoolers' corporeal synchronization with music. In M. Baroni, A. R. Addessi, R. Caterina, & M. Costa (Eds.), *Proceedings of the 9th International Conference on Music Perception and Cognition, Bologna, Italy* (pp. 472–476). Adelaide: Causal Productions.

Eerola, T., & Vuoskoski, J. K. (2013). A review of music and emotion studies: Approaches, emotion models, and stimuli. *Music Perception*, *30*, 307–340.

Egermann, H., Fernando, N., Chuen, L., & McAdams, S. (2015). Music induces universal emotion-related psychophysiological responses: Comparing Canadian listeners to Congolese Pygmies. *Frontiers in Psychology*, *5*, 1341.

Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., & Taub, E. (1995, October 13). Increased cortical representation of the fingers of the left hand in string players. *Science*, *270*, 305–307.

Eldar, E., Ganor, O., Admon, R., Bleich, A., & Hendler, T. (2007). Feeling the real world: Limbic response to music depends on related content. *Cerebral Cortex*, *17*, 2828–2840.

Ellis, R. J., Norton, A. C., Overy, K., Winner, E., Alsop, D. C., & Schlaug, G. (2012). Differentiating maturational and training influences on fMRI activation during music processing. *NeuroImage*, *60*, 1902–1912.

Epstein, D. (1995). *Shaping time: Music, the brain, and performance*. New York: Schirmer.

Ericsson, K. A. (Ed.). (1996). *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games*. Mahwah, NJ: Lawrence Erlbaum Associates.

Ericsson, K. A. (2007). Deliberate practice and the modifiability of body and mind: Toward a science of the structure and acquisition of expert and elite performance. *International Journal of Sport Psychology*, *38*, 4–34.

Ericsson, K. A. (2014). Why expert performance is special and cannot be extrapolated from studies of performance in the general population: A response to criticisms. *Intelligence*, *45*, 81–103.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*, 363–406.

Exner, S. (1875). Experimentelle Untersuchung der einfachsten psychischen Processe [Experimental study of the most simple psychological processes]. *Pflügers Archiv für die gesamte Physiologie*, *11*, 403–432.

Eysenck, M. W. (1992). *Anxiety: The cognitive perspective.* Hove: Lawrence Erlbaum.

Falconer, T. (2016). *Bad singer*. Toronto: Anasi Press.

Farnsworth, P. R. (1958). *The social psychology of music*. Austin, TX: Holt, Rinehart & Winston.

Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development*, *60*, 1497–1510.

Fifer, W. P., & Moon, C. (1989). Psychobiology of newborn auditory preferences. *Seminars in Perinatology*, *13*, 430–433.
Finney, S. A. (1997). Auditory feedback and musical keyboard performance. *Music Perception*, *15*, 153–174.
Finney, S. A., & Palmer, C. (2003). Auditory feedback and memory for music performance: Sound evidence for an encoding effect. *Memory & Cognition*, *31*, 51–64.
Fischer, S., Hallschmid, M., Elsner, A. L., & Born, J. (2002). Sleep forms memory for finger skills. *Proceedings of the National Academy of Sciences of the United States of America*, *99*, 11987–11991.
Fischinger, T., Frieler, K., & Louhivuori, J. (2015). Influence of virtual room acoustics on choir singing. *Psychomusicology: Music, Mind, & Brain*, *25*, 208–218.
Fitzpatrick, P., Schmidt, R. C., & Lockman, J. J. (1996). Dynamical patterns in the development of clapping. *Child Development*, *67*, 2691–2708.
Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press.
Fogel, H. (2007, December 11 & 12). Formality in concert: Have we gone too far? [Blog post]. Retrieved June 20, 2009 from www.artsjournal.com/ontherecord
Folkestad, G. (1996). *Computer based creative music making: Young people's music in the digital age*. Gothenburg: Acta Universitatis Gothoburgensis.
Ford, L., & Davidson, J. W. (2003). An investigation of members' roles in wind quintets. *Psychology of Music*, *31*, 53–74.
Forsyth, M. (1985). *Buildings for music*. Cambridge, MA: MIT Press.
Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *Psychology of music* (pp. 149–181). San Diego, CA: Academic Press.
Francès, R. (1958/1988). *The perception of music* (W. J. Dowling, Trans.). Hillsdale, NJ: Lawrence Erlbaum.
Frankland, B. W., & Cohen, A. J. (2004). Parsing of melody: Quantification and testing of the local grouping rules of Lerdahl and Jackendoff's *A generative theory of tonal music*. *Music Perception*, *21*, 499–543.
Friberg, A., & Sundberg, J. (1999). Does music performance allude to locomotion? A model of final *ritardandi*, derived from measurements of stopping runners. *Journal of the Acoustical Society of America*, *105*, 1469–1484.
Friberg, A., & Sundström, A. (2002). Swing ratios and ensemble timing in jazz performance: Evidence for a common rhythmic pattern. *Music Perception*, *19*, 333–349.
Fritz, T. (2012). The dock-in model of music culture and cross-cultural perception. *Music Perception*, *30*, 511–516.
Fritz, T., Ciupek, M., Kirkland, A., Ihme, K., Guha, A., Hoyer, J., & Villiringer, A. (2014). Enhanced response to music in pregnancy. *Psychophysiology*, *51*, 905–911.
Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., … Koelsch, S. (2009). Universal recognition of three basic emotions in music. *Current Biology*, *19*, 573–576.
Fuelberth, R. J. V. (2003). The effect of left hand conducting gesture on inappropriate vocal tension in individual singers. *Bulletin of the Council for Research in Music Education*, *157*, 62–70.
Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2009). Beta and gamma rhythms in human auditory cortex during musical beat processing. *Annals of the New York Academy of Sciences*, *1169*, 89–92.
Fujiyama, S., Ema, K., & Iwamiya, S. (2013). Effect of technique of conflict between music and moving picture employed in a movie directed by Akira Kurosawa. *Journal of the Acoustical Society of Japan*, *69*, 387–396. [Published in Japanese]
Gabrielsson, A. (1999). The performance of music. In D. Deutsch (Ed.), *The psychology of music* (pp. 501–602). San Diego, CA: Academic Press.
Gabrielsson, A. (2011). *Strong experiences with music* (R. Bradbury, Trans.). Oxford: Oxford University Press.

Gabrielsson, A., & Lindström, E. (2010). *The role of structure in the musical expression of emotions*. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion* (pp. 367–400). Oxford: Oxford University Press.

Gade, A. C. (2007). Acoustics in halls for speech and music. In T. D. Rossing (Ed.), *Springer handbook of acoustics* (pp. 301–350). New York: Springer.

Gade, A. C. S. (2015). Classical musicians' perception of room acoustic conditions. *Psychomusicology: Music, Mind, & Brain*, *25*, 232–235.

Gagnon, L., & Peretz, I. (2003). Mode and tempo relative contributions to 'happy–sad' judgements in equitone melodies. *Cognition & Emotion*, *17*, 25–40.

Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review*, *13*, 361–377.

Gallese, V., Gernsbacher, M. S., Heyes, C. Hickok, G., & Iacoboni, M. (2011). Mirror neuron forum. *Perspectives on Psychological Science*, *6*, 369–407.

Galton, F., Sir (1979). *Hereditary genius: An inquiry into its laws and consequences*. London: Julian Friedman. (Original work published 1869)

Garrido, S., & Schubert, E. (2011). Individual differences in the enjoyment of negative emotion in music: A literature review and experiment. *Music Perception*, *28*, 279–295.

Garrido, S., & Schubert, E. (2015). Moody melodies: Do they cheer us up? *Psychology of Music*, *43*, 244–261.

Gaser, C., & Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *Journal of Neuroscience*, *23*, 9240–9245.

Gaunt, H. (2008). One-to-one tuition in a conservatoire: The perceptions of instrumental and vocal teachers. *Psychology of Music*, *36*, 215–245.

Gentner, D. R. (1987). Timing of skilled motor performance: Tests of the proportional duration model. *Psychological Review*, *94*, 255–276.

Gerhardt, K. J., & Abrams, R. M. (2000). Fetal exposures to sound and vibroacoustic stimulation. *Journal of Perinatology*, *20*, S20–S29.

Geringer, J. M., MacLeod, R. B., Madsen, C. K., & Napoles, J. (2015). Perception of melodic intonation in performances with and without vibrato. *Psychology of Music*, *43*, 675–685.

Gerry, D. W., Faux, A. L., & Trainor, L. J. (2010). Effects of Kindermusik training on infants' rhythmic enculturation. *Developmental Science*, *13*, 545–551.

Gerry, D. W., Unrau, A., & Trainor, L. J. (2012). Active music classes in infancy enhance musical, communicative, and social development. *Developmental Science*, *15*, 398–407.

Gershon, R. R. M., Neitzel, R., Barrera, M. A., & Akram, M. (2006). Pilot survey of subway and bus stop noise levels. *Journal of Urban Health: Bulletin of the New York Academy of Medicine*, *83*, 802–812.

Gerson, S. A., Schiavio, A., Timmers, R., & Hunnius, S. (2015). Active drumming experience increases infants' sensitivity to audiovisual synchrony during observed drumming actions. *PLOS ONE*, *10*, e0130960.

Getz, L. M., Marks, S., & Roy, M. (2014). The influence of stress, optimism, and music training on music uses and preferences. *Psychology of Music*, *42*, 71–85.

Gfeller, K., Christ, A., Knutsen, J. F., Witt, S., Murray, K. T., & Tyler, R. S. (2000). Musical backgrounds, listening habits, and aesthetic enjoyment of adult cochlear implant recipients. *Journal of the American Academy of Audiology*, *11*, 390–406.

Gfeller, K., Witt, S., Woodworth, G., Mehr, M. A., & Knutsen, J. F. (2002). Effects of frequency, instrumental family, and cochlear implant type on timbre recognition and appraisal. *Annals of Otology, Rhinology, and Laryngology*, *111*, 349–356.

Ginsborg, J. (2002). Classical singers learning and memorizing a new song: An observational study. *Psychology of Music*, *30*, 58–101.

Ginsborg, J., & Sloboda, J. A. (2007). Singers' recall for the words and melody of a new, unaccompanied song. *Psychology of Music*, *35*, 421–440.

Gobet, F., & Campitelli, G. (2007). The role of domain-specific practice, handedness, and starting age in chess. *Developmental Psychology*, *43*, 159–172.

Goebl, W. (2001). Melody lead in piano performance: Expressive device or artifact? *Journal of the Acoustical Society of America*, *110*, 563–572.

Goebl, W., & Palmer, C. (2009). Synchronization of timing and motion among performing musicians. *Music Perception*, *26*, 427–438.

Goldberg, P. (1968). Are women prejudiced against women? *Transaction*, *5*, 316–322.

Goldin, C., & Rouse, C. (2000). Orchestrating impartiality: The impact of 'blind' auditions on female musicians. *American Economic Review*, *90*, 715–741.

Goldman, A. (2013). Towards a cognitive–scientific research program for improvisation: Theory and an experiment. *Psychomusicology: Music, Mind and Brain*, *23*, 210–221.

Gomez, P., & Danuser, B. (2007). Relationships between musical structure and psychophysiological measures of emotion. *Emotion*, *7*, 377–387.

Good, A., & Russo, F. A. (2016). Singing promotes cooperation in a diverse group of children. *Social Psychology*, *47*, 340–344.

Gorbman, C. (1987). *Unheard melodies: Narrative film music*. Bloomington, IN: Indiana University Press.

Gordon, R. L., Jacobs, M. S. Schuele, C. M., & McAuley, J. D. (2015). Perspectives on the rhythm–grammar link and its implications for typical and atypical language development. *Annals of the New York Academy of Sciences*, *1337*, 16–25.

Gordon, R. L., Shivers, C. M., Wieland, E. A., Kotz, S. A., Yoder, P. J., & Devin McAuley, J. (2015). Musical rhythm discrimination explains individual differences in grammar skills in children. *Developmental Science*, *18*, 635–644.

Gosselin, N., Peretz, I., Noulhiane, M., Hasboun, D., Beckett, C., Baulac, M., & Samson, S. (2005). Impaired recognition of scary music following unilateral temporal lobe excision. *Brain*, *128*, 628–640.

Grahn, J. A. (2012). Neural mechanisms of rhythm perception: Current findings and future perspectives. *Topics in Cognitive Science*, *4*, 585–606.

Grahn, J. A., & Rowe, J. B. (2009). Feeling the beat: Premotor and striatal interactions in musicians and nonmusicians during beat perception. *Journal of Neuroscience*, *29*, 7540–7548.

Granier-Deferre, C., Bassereau, S., Ribeiro, A., Jacquet, A.-Y., & DeCasper, A. J. (2011). A melodic contour repeatedly experienced by human near-term fetuses elicits a profound cardiac reaction one month after birth. *PLOS ONE*, *6*, e17304.

Gratier, M., Devouche, E., Guellai, B., Infanti, R., Yilmaz, E., & Parlato-Oliveira, E. (2015). Early development of turn-taking in vocal interaction between mothers and infants. *Frontiers in Psychology*, *6*, 1167.

Green, G. A. (1990). The effect of vocal modeling on pitch-matching accuracy of elementary schoolchildren. *Journal of Research in Music Education*, *38*, 225–331.

Grewe, O., Nagel, F., Kopiez, R., & Altenmüller, E. (2007). Listening to music as a re-creative process: Physiological, psychological, and psychoacoustical correlates of chills and strong emotions. *Music Perception*, *24*, 297–314.

Griffiths, T. D., Johnsrude, I., Dean, J. L., & Green, G. G. R. (1999). A common neural substrate for the analysis of pitch and duration pattern in segmented sound? *NeuroReport*, *10*, 3825–3830.

Gruson, L. M. (1988). Rehearsal skill and musical competence: Does practice make perfect? In J. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition* (pp. 91–112). Oxford: Oxford University Press.

Guhn, M., Hamm, A., & Zentner, M. (2007). Physiological and musico-acoustic correlates of the chill response. *Music Perception*, *24*, 473–483.

Hajda, J. M., Kendall, R. A., Carterette, E. C., & Harshberger, M. L. (1997). Methodological issues in timbre research. In I. Deliège & J. Sloboda (Eds.), *Perception and cognition of music* (pp. 253–306). Hove: Psychology Press.

Hall, D. E. (2001). *Musical acoustics*. Pacific Grove, CA: Brooks/Cole.

Hallam, S. (1995). Professional musicians' approaches to the learning and interpretation of music. *Psychology of Music*, *23*, 111–128.

Hallam, S. (1997). Approaches to instrumental music practice of experts and novices: Implications for education. In H. Jørgensen & A. C. Lehmann (Eds.), *Does practice make perfect? Current theory and research on instrumental music practice* (pp. 89–107). Oslo: Norges Musikkhøgskole.

Hallam, S., Rinta, T., Varvarigou, M., Creech, A., Papageorgi, I., Gomes, T., & Lanipekun, J. (2012). The development of practising strategies in young people. *Psychology of Music*, *40*, 652–680.

Hallett, R., & Lamont, A. (2016, October 25). Evaluation of a motivational pre-exercise music intervention. *Journal of Health Psychology*. pii: 1359105316674267. [Epub ahead of print]

Halpern, A. R. (1989). Memory for the absolute pitch of familiar songs. *Memory & Cognition*, *17*, 572–581

Halpern, A. R. (2003). Cerebral substrates of musical imagery. In I. Peretz & R. Zatorre (Eds.), *The cognitive neuroscience of music* (pp. 217–230). New York: Oxford University Press.

Halpern, A. R., & Bartlett, J. C. (2010). Memory for melodies. In M. R. Jones, R. R. Fay, & A. N. Popper (Eds.), *Music Perception* (pp. 233–258). New York: Springer.

Halpern, A. R., & Bartlett, J. C. (2011). The persistence of musical memories: A descriptive study of earworms. *Music Perception*, *28*, 425–432.

Halwani, G. F., Loui, P., Rueber, T., & Schlaug, G. (2011). Effects of practice and experience on the arcuate fasciculus: Comparing singers, instrumentalists, and non-musicians. *Frontiers in Psychology*, *2*, 156.

Hambrick, D. Z., & Meinz, E. J. (2011). Limits on the predictive power of domain-specific experience and knowledge in skilled performance. *Current Directions in Psychological Science*, *20*, 275–279.

Hambrick, D. Z., Oswald, F. L., Altmann, E. M., Meinz, E. J., Gobet, F., & Campitelli, G. (2014). Deliberate practice: Is that all it takes to become an expert? *Intelligence*, *45*, 34–45.

Hambrick, D. Z., & Tucker-Drob, E. M. (2015). The genetics of music accomplishment: Evidence for gene–environment correlation and interaction. *Psychonomic Bulletin & Review*, *22*, 112–120.

Handel, S. (1989). *Listening: An introduction to the perception of auditory events*. Cambridge, MA: MIT Press.

Hannon, E. E., Soley, G., & Ullal, S. (2012). Familiarity overrides simplicity in rhythmic pattern perception: A cross-cultural examination of American and Turkish listeners. *Journal of Experimental Psychology: Human Perception and Performance*, *38*, 543–548.

Hannon, E. E., & Trehub, S. E. (2005a). Metrical categories in infancy and adulthood. *Psychological Science*, *16*, 48–55.

Hannon, E. E., & Trehub, S. E. (2005b). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences*, *102*, 12639–12643.

Hanslick, E. (1986). *On the musically beautiful* (G. Payzant, Trans.). Indianapolis, IN: Hackett. (Original work published 1891)

Hargreaves, D. J., & North, A. C. (Eds.). (1997). *The social psychology of music*. Oxford: Oxford University Press.

Harré, R. (1997). Emotion in music. In M. Hjort & S. Laver (Eds.), *Music and the arts* (pp. 110–118). New York: Oxford University Press.

Harrison, A. C., & O'Neill, S. A. (2000). Children's gender-typed preferences for musical instruments: An intervention study. *Psychology of Music*, *28*, 81–97.

Harrison, G. (1980). *I, me, mine*. New York: Simon & Schuster.

Hasegawa, A., Okanoya, K., Hasegawa, T., & Seki, Y. (2011). Rhythmic synchronization tapping to an audio-visual metronome in budgerigars. *Scientific Reports*, *1*, 120.

Harrop-Allin, S. (2010). Recruiting learners' musical games as resources for South African music education, using a multiliteracies approach (Doctoral thesis, University of the Witwatersrand, Johannesburg, South Africa). Retrieved from http://wiredspace.wits.ac.za/handle/10539/8894

Hatfield, J. L. (2016). Performing at the top of one's musical game. *Frontiers in Psychology*, *7*, 1356.

Hattori, Y., Tomonaga, M., & Matsuzawa, T. (2013). Spontaneous synchronized tapping to an auditory rhythm in a chimpanzee. *Scientific Reports*, *3*, 1566.

Hayes, J. R. (1981). *The complete problem solver*. Philadelphia, PA: Franklin Institute Press.

He, C., Hotson, L., & Trainor. L. J. (2007). Mismatch responses to pitch changes in early infancy. *Journal of Cognitive Neuroscience*, *19*, 878–892.
He, C., Hotson, L., & Trainor, L. J. (2009). Maturation of cortical mismatch responses to occasional pitch change in early infancy: Effects of presentation rate and magnitude of change. *Neuropsychologia*, *47*, 218–229.
Hellman, R. P. (1976). Growth of loudness at 1000 and 3000 Hz. *Journal of the Acoustical Society of America*, *60*, 672–679.
Helmholtz, H. L. F. von. (1863). *On the sensations of tone as the physiological basis of the theory of music*. New York: Dover.
Hepper, P. G. (1991). An examination of fetal learning before and after birth. *Irish Journal of Psychology*, *12*, 95–107.
Hepper, P. G., & Shahidullah, B. S. (1994). Development of fetal hearing. *Archives of Disease in Childhood*, *71*, F81–F87.
Hevner, K. (1936). Experimental studies of the elements of expression in music. *American Journal of Psychology*, *48*, 246–268.
Highben, Z., & Palmer, C. (2004). Effects of auditory and motor mental practice in memorized piano performance. *Bulletin of the Council for Research in Music Education*, *159*, 58–65.
Hirsch, I. J. (1959). Auditory perception of temporal order. *Journal of the Acoustical Society of America*, *31*, 759–767.
Hodgetts, W. E, Rieger, J. M., & Szarko, R. A. (2007). The effects of listening environment and earphone style on preferred listening levels of normal hearing adults using an MP3 player. *Ear and Hearing*, *28*, 290–297.
Hoeckner, B., Wyatt, E. W., Decety, J., & Nusbaum, H. (2011). Film music influences how viewers relate to movie characters. *Psychology of Aesthetics, Creativity, and the Arts*, *5*, 146–153.
Hoeschele, M., Weisman, R. G., Guillette, L. M., Hahn, A. H., & Sturdy, C. B. (2013). Chickadees fail standardized operant tests for octave equivalence. *Animal Cognition*, *16*, 599–609.
Hofman, P. M., van Riswick, J. G. A., & van Opstal, A. J. (1998). Relearning sound localization with new ears. *Nature Neuroscience*, *1*, 417–421.
Honing, H. (2005). Is there a perception-based alternative to kinematic models of tempo rubato? *Music Perception*, *23*, 79–85.
Honing, H. (2013). Structure and interpretation of rhythm in music. In D. Deutsch (Ed.), *The psychology of music* (3rd ed., pp. 369–404). Amsterdam: Academic Press.
Honing, H., Merchant, H., Háden, G. P., Prado, L., & Bartolo, R. (2012). Rhesus monkeys (*Macaca mulatta*) detect rhythmic groups in music, but not the beat. *PLOS ONE*, *7*, e51369.
Hood, M. (1966). *Sléndro* and *pélog* redefined. *Selected Reports, UCLA Institute of Ethnomusicology*, *1*, 28–37.
Hoover, A., & Krishnamurti, S. (2010). Survey of college students' MP3 listening: Habits, safety issues, attitudes, and education. *American Journal of Audiology*, *19*, 73–83.
Hoppe, D., Sadakata, M., & Desain, P. (2006). Development of real-time visual feedback assistance in singing training: A review. *Journal of Computer Assisted Learning*, *22*, 308–316.
Hopyan, T., Manno, F. A. M., Papsin, B. C., & Gordon, K. A. (2016). Sad and happy emotion discrimination in music by children with cochlear implants. *Child Neuropsychology*, *22*, 366–380.
Hospers. J. (1964). *Meaning and truth in the arts*. Hamden, CT: Archon Books.
Hove, M. J., & Risen, J. L. (2009). It's all in the timing: Interpersonal synchrony increases affiliation. *Social Cognition*, *27*, 949–960.
Howe, M. J. A., Davidson, J. W., & Sloboda, J. A. (1998). Innate talents: Reality or myth? *Behavioral and Brain Sciences*, *21*, 399–442.
Howell, P., Powell, D. J., & Khan, I. (1983). Amplitude contour of the delayed signal and interference in delayed auditory feedback tasks. *Journal of Experimental Psychology: Human Perception and Performance*, *9*, 772–784.
Hulse, S. H., Cynx, J., & Humpal, J. (1984). Absolute and relative pitch discrimination in serial pitch perception by birds. *Journal of Experimental Psychology: General*, *113*, 38–54.

Hunt, R. R., & McDaniel, M. A. (1993). The enigma of organization and distinctiveness. *Journal of Memory & Language*, *32*, 421–445.
Hunter, P. G., Schellenberg, E. G., & Schimmack, U. (2010). Feelings and perceptions of happiness and sadness induced by music: Similarities, differences, and mixed emotions. *Psychology of Aesthetics, Creativity, and the Arts*, *4*, 47–56.
Hunter, P. G., Schellenberg, E. G., & Stalinski, S. M. (2011). Liking and identifying emotionally expressive music: Age and gender differences. *Journal of Experimental Child Psychology*, *110*, 80–93.
Huron, D. (2003). Is music an evolutionary adaptation? In I. Peretz & R. Zatorre (Eds.), *The cognitive neuroscience of music* (pp. 57–78). New York: Oxford University Press.
Huron, D. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge, MA: MIT Press.
Huron, D. (2011). Why is sad music pleasurable? A possible role for prolactin. *Musicae Scientiae*, *15*, 146–158. [Special issue on music and emotion]
Husain, G., Thompson, W. F., & Schellenberg, E. G. (2002). Effects of musical tempo and mode on arousal, mood, and spatial abilities. *Music Perception*, *20*, 151–171.
Hutchins, S., & Palmer, C. (2008). Repetition priming in music. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 693–707.
Hutchins, S., & Peretz, I. (2012). A frog in your throat or in your ear? Searching for the causes of poor singing. *Journal of Experimental Psychology: General*, *141*, 76–97.
Hutchins, S., & Peretz, I. (2013). Vocal pitch shift in congenital amusia (pitch deafness). *Brain & Language*, *125*, 106–117.
Hutchins, S., Zarate, J. M., Zatorre, R. J., & Peretz, I. (2010). An acoustical study of vocal pitch matching in congenital amusia. *Journal of the Acoustical Society of America*, *127*, 504–512.
Hyde, K. L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A. C., & Schlaug, G. (2009). Musical training shapes structural brain development. *Journal of Neuroscience*, *29*, 3019–3025.
Hyde, K. L., Lerch, J. P., Zatorre, R. J., Griffiths, T. D., Evans, A. C., & Peretz, I. (2007). Cortical thickness in congenital amusia: When less is better than more. *Journal of Neuroscience*, *27*, 13028–13032.
Hyde, K. L., Zatorre, R. J., Griffiths, T. D., Lerch, J. P., & Peretz, I. (2006). Morphometry of the amusic brain: A two-site study. *Brain*, *129*, 2562–2570.
Ilari, B. (2015). Rhythmic engagement with music in early childhood: A replication and extension. *Journal of Research in Music Education*, *62*, 332–343.
Ilari, B., & Polka, L. (2006). Music cognition in early infancy: Infants' preferences and long-term memory for Ravel. *International Journal of Music Education*, *24*, 7–20.
Ingram, J. (1997, May 11). More proof music lifts young IQs. *The Toronto Star*, p. F8.
Innes-Brown, H., Marozeau, J. P., Storey, C. M., & Blamey, P. J. (2013). Tone, rhythm, and timbre perception in school-age children using cochlear implants and hearing aids. *Journal of the American Academy of Audiology*, *24*, 789–806.
Ivanov, V. K., & Geake, J. G. (2003). The Mozart effect and primary school children. *Psychology of Music*, *31*, 405–413.
Iyer, V. (2002). Embodied mind, situated cognition, and expressive microtiming in African-American music. *Music Perception*, *19*, 387–414.
Jabusch, H.-C., & Altenmüller, E. (2006a). Epidemiology, phenomenology, and therapy of musician's cramp. In E. Altenmüller, M. Wisendanger, & J. Kesselring (Eds.), *Music, motor control and the brain* (pp. 265–282). Oxford: Oxford University Press.
Jabusch, H.-C., & Altenmüller, E. (2006b). Focal dystonia in musicians: From phenomenology to therapy. *Advanced Cognitive Psychology*, *2*, 207–220.
Jacob, C., Guéguen, N., Boulbry, G., & Sami, S. (2009). 'Love is in the air': Congruence between background music and goods in a flower shop. *International Review of Retail, Distribution and Consumer Research*, *19*, 75–79.
James, W. (1890). *Principles of psychology* (Vol. 1). New York: Holt.

Janata, P. (2009). The neural architecture of music-evoked autobiographical memories. *Cerebral Cortex*, *19*, 2579–2594.

Janata, P., Birk, J., van Horn, J., Leman, M., Tillmann, B., & Bharucha, J. (2002). The cortical topography of tonal structures underlying Western music. *Science*, *298*, 2167–2170.

Janata, P., Tomic, S. T., & Haberman, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. *Journal of Experimental Psychology: General*, *141*, 54–75.

Jardri, R., Pins, D., Houfflin-Debarge, V., Chaffiotte, C., Rocourt, N., Pruvo, J.-P., … Thomas, P. (2008). Fetal cortical activation to sound at 33 weeks of gestation: A functional MRI study. *NeuroImage*, *42*, 10–18.

Jiang, W., Zhao, F., Guderley, N., & Manchaiah, V. (2016). Daily music exposure dose and hearing problems using personal listening devices in adolescents and young adults: A systematic review. *International Journal of Audiology*, *55*, 197–205.

Johansen, G. G. (2017). Explorational experimental practice. *Psychology of Music*. DOI: 10.1177/0305735617695657

Johnson-Laird, P. N. (2002). How jazz musicians improvise. *Music Perception*, *19*, 415–442.

Jones, A. M. (1959). *Studies in African music*. London: Oxford.

Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, *83*, 323–355.

Jones, M. R. (1987). Dynamic pattern structure in music: Recent theory and research. *Perception & Psychophysics*, *41*, 621–634.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, *96*, 459–491.

Jones, M. R., Moynihan, H. M., MacKenzie, N., & Puente, J. K. (2002). Stimulus-driven attending in dynamic auditory arrays. *Psychological Science*, *13*, 313–319.

Joseph, R. (2000). Fetal brain behavior and cognitive development. *Developmental Review*, *20*, 81–98.

Jusczyk, P. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press.

Jusczyk, P. W., & Krumhansl, C. L. (1993). Pitch and rhythmic patterns affecting infants' sensitivity to musical phrase structure. *Journal of Experimental Psychology: Human Perception & Performance*, *19*, 627–640.

Juslin, P. N. (1997). Emotion communicated in music performance: A functionalist perspective and some data. *Music Perception*, *14*, 383–418.

Juslin, P. N. (2000). Cue utilization in communication of emotion in music performance: Relating performance to perception. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 1797–1813.

Juslin, P. N. (2013a). From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews*, *10*, 235–266.

Juslin, P. N. (2013b). What does music express? Basic emotions and beyond. *Frontiers in Psychology*, *4*, 596.

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, *129*, 770–814.

Juslin, P. N., Liljeström, S., Laukka, P., Västfjäll, D., & Lundqvist, L.-O. (2011). Emotional reactions to music in a nationally representative sample of Swedish adults: Prevalence and causal influences. *Musicae Scientiae*, *15*, 174–207.

Juslin, P. N., & Sloboda, J. A. (2001). *Music and emotion: Theory and research*. Oxford: Oxford University Press.

Juslin, P. N., & Sloboda, J. A. (2010). *Handbook of music and emotion*. Oxford: Oxford University Press.

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, *31*, 559–621.

Justus, T. C., & Bharucha, J. (2001). Modularity in musical processing: The automaticity of harmonic priming. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 1000–1011.

Justus, T., & Hutsler, J. J. (2005). Fundamental issues in the evolutionary psychology of music: Assessing innateness and domain specificity. *Music Perception*, *23*, 1–28.

Kaas, J. H., & Hackett, T. A. (2005). Subdivisions and connections of the auditory cortex in primates: A working model. In R. König, P. Heil, E. Budinger, & H. Scheich (Eds.), *The auditory cortex: A synthesis of human and animal research* (pp. 7–25). Mahwah, NJ: Lawrence Erlbaum.

Kammler, D. (2008). *A first course in Fourier analysis*. Cambridge: Cambridge University Press.

Kastner, M. P., & Crowder, R. G. (1990). Perception of the major/minor distinction: IV. Emotional connotations in young children. *Music perception*, *8*, 189–202.

Kawakami, A., Furukawa, K., Katahira, K., & Okanoya, K. (2013). Sad music induces pleasant emotion. *Frontiers in Psychology*, *4*, 311.

Kawakami, A., Furukawa, K., Katahira, K., Kamiyama, K., & Okanoya, K. (2013). Relations between musical structures and perceived and felt emotions. *Music Perception*, *30*, 407–417.

Keefe, D. H., Bulen, J. C., Campbell, S. L., & Burns, E. M. (1994). Pressure transfer function and absorption cross section from the diffuse field to the human infant ear canal. *Journal of the Acoustical Society of America*, *95*, 355–371.

Keeler, J. R., Roth, E. A., Neuser, B. L., Spitsbergen, J. M., Waters, D. J. M., & Vianney, J.-M. (2015). The neurochemistry and social flow of singing: bonding and oxytocin. *Frontiers in Human Neuroscience*, *9*, 518.

Keller, P. E. (2001). Attentional resource allocation in musical ensemble performance. *Psychology of Music*, *29*, 20–38.

Kemp, A. E. (1996). *The musical temperament: Psychology and personality of musicians*. Oxford: Oxford University Press.

Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science*, *11*, 201–258.

Kendall, R. A. (1986). The role of acoustic signal partitions in listener categorization of musical phrases. *Music Perception*, *4*, 185–214.

Kenny, D. (2011). *The psychology of music performance anxiety*. Oxford: Oxford University Press.

Kenny, D. T., Davis, P., & Oates, J. (2004). Music performance anxiety and occupational stress amongst opera chorus artists and their relationship with state and trait anxiety and perfectionism. *Journal of Anxiety Disorders*, *18*, 757–777.

Kerker, G. (1904). Opinions of some New York leaders on women as orchestral players. *Musical Standard*, *21*, 217–218.

Kessler, E. J., Hansen, C., & Shepard, R. N. (1984). Tonal schemata in the perception of music in Bali and in the West. *Music Perception*, *2*, 131–165.

Kieslowski, K. (Director), & Karmitz, M. (Producer). (1993). *Three colors: Blue* [Motion picture]. Paris: CAB Productions. (Available from Miramax Films)

Kim, J.-Y., Kim, M.-S., Min, S.-N., Cho, Y.-J., & Choi, J. (2012). Prevalence of playing-related musculoskeletal disorders in traditional Korean string instrument players. *Medical Problems of Performing Artists*, *27*, 212–218c.

Kirschner, S., & Tomasello, M. (2009). Joint drumming: Social context facilitates synchronization in preschool children. *Journal of Experimental Child Psychology*, *102*, 299–314.

Kirschner, S., & Tomasello, M. (2010). Joint music making promotes prosocial behavior in 4-year-old children. *Evolution and Human Behavior*, *31*, 354–364.

Kisilevsky, B. S., Hains, S. M. J., Jacquet, A.-Y., Granier-Deferre, C., & Lecanuet, J. P. (2004). Maturation of fetal responses to music. *Developmental Science*, *7*, 550–559.

Kivy, P. (1990). *Music alone*. Ithaca, NY: Cornell University Press.

Kleber, B., Zeitouni, A. G., Friberg, A., & Zatorre, R. J. (2013). Experience-dependent modulation of feedback integration during singing: Role of the right anterior insula. *Journal of Neuroscience*, *33*, 6070–6080.

Knight, T., Upham, F., & Fujinaga, I. (2011). The potential for automatic assessment of trumpet tone quality. In *Proceedings of the 12th International Society for Music Information Retrieval Conference* (pp. 573–578). Miami, FL: International Society for Music Information Retrieval.

Knöferle, K. M., Spangenberg, E. R., Herrmann, A., & Landwehr, J. R. (2012). It's all in the mix: The interactive effect of music tempo and mode on in-store sales. *Marketing Letters*, *23*, 325–337.

Knox, C. A. H. (2009). *The effect of visual feedback of sound intensity on preferred iPod listening levels* (Doctoral dissertation). Retrieved from ProQuest Information & Learning (Order No. AAI3352761).

Knox, R. A. (1993, October 14). Mozart makes you smarter, Calif. researchers suggest. *The Boston Globe*, p. 1.

Knudsen, E. I., & Knudsen, P. F. (1989). Vision calibrates sound localization in developing barn owl. *Journal of Neuroscience*, *9*, 3306–3313.

Koelsch, S., Gunter, T. C., von Cramon, D. Y., Zysset, S., Lohmann, G., & Frederici, A. D. (2002). Bach speaks: A cortical 'language-network' serves the processing of music. *NeuroImage*, *17*, 956–966.

Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., & Frederici, A. D. (2004). Music, language and meaning: Brain signatures of semantic processing. *Nature Neuroscience*, *7*, 302–307.

Kohlberg, G. D., Mancuso, D. M., Chari, D. A., & Lalwani, A. K. (2015). Music engineering as a novel strategy for enhancing music enjoyment in the cochlear implant recipient. *Behavioural Neurology*, *2015*, 1–7.

Kohn, D., & Eitan, Z. (2016). Moving music: Correspondences of musical parameters and movement dimensions in children's motion and verbal responses. *Music Perception*, *34*, 40–55.

Konečni, V. (2008). Does music induce emotion? A theoretical and methodological analysis. *Psychology of Aesthetics, Creativity and the Arts*, *2*, 115–129.

Koops, L. H. (2012). 'Now can I watch my video?': Exploring musical play through video sharing and social networking in an early childhood music class. *Research Studies in Music Education*, *34*, 15–28.

Kopiez, R., & Lee, J. I. (2006). Towards a dynamic model of skills involved in sight reading music. *Music Education Research*, *8*, 97–120.

Kopiez, R., Platz, F, Müller S., & Wolf, A. (2015). When the pulse of the song goes on: Fade-out in popular music and the pulse continuity phenomenon. *Psychology of Music*, *43*, 395–374.

Koskoff, E. (2014). *A feminist ethnomusicology.* Chicago, IL: University of Illinois Press.

Kozbelt, A. (2012). Process, self-evaluation, and lifespan creativity trajectory in eminent composers. In D. Collins (Ed.), *The act of musical composition* (pp. 27–52). Farnham: Ashgate.

Krampe, R. T., & Ericsson, K. A. (1995). Deliberate practice and elite musical performance. In J. Rink (Ed.), *The practice of performance: Studies in musical interpretation* (pp. 84–102). Cambridge: Cambridge University Press.

Kratus, J. (1989). A time analysis of the compositional processes used by children ages 7 to 11. *Journal of Research in Music Education*, *37*, 5–20.

Krueger, C., Holditch-Davis, D., Quint, S., & DeCasper, A. (2004). Recurring auditory experience in the 28- to 34-week-old fetus. *Infant Behavior & Development*, *27*, 537–543.

Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology*, *11*, 346–374.

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. Oxford: Oxford University Press.

Krumhansl, C. L. (1995). Music psychology and music theory: Problems and prospects. *Music Theory Spectrum*, *17*, 53–80.

Krumhansl, C. L. (1996). A perceptual analysis of Mozart's Piano Sonata K. 282: Segmentation, tension, and musical ideas. *Music Perception*, *13*, 401–432.

Krumhansl, C. L. (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology*, *51*, 336–353.

Krumhansl, C. L., & Cuddy, L. (2010). A theory of tonal hierarchies in music. In M. R. Jones, R. R. Fay & A. N. Popper (Eds.), *Music perception* (pp. 51–87). New York: Springer.

Krumhansl, C. L., & Shepard, R. N. (1979). Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 579–594.

Krumhansl, C. L., Toivanen, P., Eerola, T., Toiviainen, P., Järvinen, T., & Louhivuori, J. (2000). Cross-cultural music cognition: Cognitive methodology applied to North Sami Yoiks. *Cognition*, *76*, 13–58.

Kubik, G. (1962). The phenomenon of inherent rhythms in East and Central African instrumental music. *African Music*, *3*, 33–42.

Kujawa, S. G., & Liberman, M. C. (2006). Acceleration of age-related hearing loss by early noise exposure: Evidence of a misspent youth. *Journal of Neuroscience*, *26*, 2115–2123.

Lahav, A. (2014). Questionable sound exposure outside of the womb: Frequency analysis of environmental noise in the neonatal intensive care unit. *Acta Paediatrica*, *104*, e14–e19.

Lahav, A., Saltzman, E., & Schlaug, G. (2007). Action representation of sound: Audiomotor recognition network while listening to newly acquired actions. *Journal of Neuroscience*, *27*, 308–314.

Lahdelma, I., & Eerola, T. (2015). Single chords convey distinct emotional qualities to both naïve and expert listeners. *Psychology of Music*, *44*, 37–54.

Lahme, A., Eibl, I., & Reichl, F.-X. (2014). Typical musculoskeletal patterns in upper string players with neck and arm problems. *Medical Problems of Performing Artists*, *29*, 241–242.

Laitinen, H. M., Toppila, E. M., Olkinuora, P. S., & Kuisma, K. (2003). Sound exposure among the Finnish National Opera personnel. *Applied Occupational Environmental Hygiene*, *18*, 177–182.

Lamont, A., & Heye, A. (2010). Mobile listening situations in everyday life: The use of MP3 players while travelling. *Musicae Scientiae*, *14*, 95–120.

Langer, S. K. (1942). *Philosophy in a new key*. Cambridge, MA: Harvard University Press.

Langer, S. K. (1953). *Feeling and form*. London: Routledge & Kegan Paul.

Large, E. W., & Gray, P. M. (2015). Spontaneous tempo and rhythmic entrainment in a bonobo (*Pan paniscus*). *Journal of Comparative Psychology*, *129*, 317–328.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, *106*, 119–159.

Large, E. W., Palmer, C., & Pollack, J. B. (1995). Reduced memory representations for music. *Cognitive Science*, *19*, 53–96.

Larsen, J. T., & Stastny, B. J. (2011). It's a bittersweet symphony: Simultaneously mixed emotional responses to music with conflicting cues. *Emotion*, *11*, 1469–1473.

Lashley, K. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior* (pp. 112–136). New York: Wiley.

Laukka, P., & Gabrielsson, A. (2000). Emotional expression in drumming performance. *Psychology of Music*, *28*, 181–189.

Launay, J., Grube, M., & Stewart, L. (2014). Dysrhythmia: A specific congenital rhythm perception deficit. *Frontiers* in *Psychology*, *5*, 18.

Le Guellec, H., Guéguen, N., Jacob, C., & Pascual, A. (2007). Cartoon music in a candy store: A field experiment. *Psychological Reports*, *100*, 1255–1258.

Leaver, R., Harris, E. C., & Palmer, K. T. (2011). Musculoskeletal pain in elite professional musicians from British symphony orchestras. *Occupational Medicine*, *61*, 549–555.

Lecanuet, J.-P., & Granier-Deferre, C. (1993). Speech stimuli in the fetal environment. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage, & J. Morton (Eds.), *Developmental neurocognition* (pp. 237–248). New York: Plenum.

Lecanuet, J.-P., Granier-Deferre, C., Jacquet, A. Y., & DeCasper, A. J. (2000). Fetal discrimination of low-pitched musical notes. *Developmental Psychobiology*, *36*, 29–39.

Lee, B. S. (1950). Effects of delayed speech feedback. *Journal of the Acoustical Society of America*, *22*, 824–826.

Lee, G. Y., & Kisilevsky, B. S. (2014). Fetuses respond to father's voice but prefer mother's voice after birth. *Developmental Psychobiology*, *56*, 1–11.

Lehmann, A. C., & Ericsson, K. A. (1996). Performance without preparation: Structure and acquisition of expert sight-reading and accompanying performance. *Psychomusicology*, *15*, 1–29.

Lempert, T., & Bauer, L. T. (1995). Mass fainting at rock concerts. *New England Journal of Medicine*, *332*, 1721.

Lerdahl, F. (1988). Cognitive constraints on compositional systems. In J. A. Sloboda (Ed.), *Generative processes in music* (pp. 231–259). Oxford: Clarendon Press.

Lerdahl, F. (2001). *Tonal Pitch Space*. New York: Oxford University Press.

Lerdahl, F. (2009). Genesis and architecture of the GTTM project. *Music Perception*, *26*, 187–194.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.

Lerdahl, F., & Jackendoff, R. (1983–1984). An overview of hierarchical structure in music. *Music Perception*, *1*, 229–252.

Lerdahl, F., & Krumhansl, C. L. (2007). Modeling tonal tension. *Music Perception*, *24*, 329–366.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levinson, J. (2004). Musical chills and other delights of music. In J. W. Davidson (Ed.), *The music practitioner* (pp. 335–351). Aldershot: Ashgate.

Levitin, D. J. (1994). Absolute memory for musical pitch: Evidence from the production of learned melodies. *Perception & Psychophysics*, *56*, 414–423.

Levitin, D. J. (2008). *The world in six songs: How the musical brain created human nature*. New York: Dutton.

Levitin, D. J., & Cook, P. R. (1996). Memory for musical tempo: Additional evidence that auditory memory is absolute. *Perception & Psychophysics*, *58*, 927–935.

Levitin, D. J., & Menon, V. (2003). Musical structure is processed in 'language' areas of the brain: A possible role for Brodmann Area 47 in temporal coherence. *NeuroImage*, *20*, 2142–2152.

Liégois-Chauvel, C., Peretz, I., Babaï, M., Laguitton, V., & Chauvel, P. (1998). Contribution of different cortical areas in the temporal lobes to music processing. *Brain*, *121*, 1853–1867.

Lim, S., & Lippman, L. G. (1991). Mental practice and memorization of piano music. *Journal of General Psychology*, *118*, 21–30.

Lima, C. F., & Castro, S. L. (2011a). Emotion recognition in music changes across the adult lifespan. *Cognition and Emotion*, *25*, 585–598.

Lima, C. F., & Castro, S. L. (2011b). Speaking to the trained ear: Musical expertise enhances the recognition of emotions in speech prosody. *Emotion*, *11*, 1021–1031.

Limb, C. J., & Braun, A. R. (2008). Neural substrates of spontaneous musical performance: An fMRI study of jazz improvisation. *PLOS ONE*, *3*, e1679.

Linklater, F. (1997). Effects of audio- and videotape models on performance achievement of beginning clarinetists. *Journal of Research in Music Education*, *45*, 402–414.

Loehr, J. D., & Palmer, C. (2007). Cognitive and biomechanical influences in pianists' finger tapping. *Experimental Brain Research*, *178*, 518–528.

Loewy, J., Stewart, K., Dassler, A. M., Telsey, A., & Homel, P. (2013). The effects of music therapy on vital signs, feeding, and sleep in premature infants. *Pediatrics*, *131*, 902–918.

London, J. (1995). Some examples of complex meters and their implications for models of metric perception. *Music Perception*, *13*, 59–77.

London, J. (2004). *Hearing in time: Psychological aspects of musical meter*. New York: Oxford University Press.

London, J. (2011). Tactus ≠ tempo: Some dissociations between attentional focus, motor behavior, and tempo judgment. *Empirical Musicology Review*, *6*, 43–55.

Longhi, E. (2009). 'Songese': Maternal structuring of musical interaction with infants. *Psychology of Music*, *37*, 195–213.

Longuet-Higgins, H. C., & Lee, C. S. (1982). The perception of musical rhythms. *Perception*, *11*, 115–128.

Lonsdale, K., Laakso, E.-L., & Tomlinson, V. (2014). Contributing factors, prevention, and management of playing-related musculoskeletal disorders among flute players internationally. *Medical Problems of Performing Artists*, *29*, 155–162.

Loui, P., Alsop, D., & Schlaug, G. (2009). Tone deafness: A new disconnection syndrome? *Journal of Neuroscience*, *29*, 10215–10220.

Loui, P., Guenther, F. H., Mathys, C., & Schlaug, G. (2008). Action-perception mismatch in tone-deafness. *Current Biology*, *18*, R331–R332.

Loui, P., Zamm, A., & Schlaug, G. (2012). Enhanced functional networks in absolute pitch. *NeuroImage*, *63*, 632–640.

Luck, G., & Sloboda, J. A. (2008). Exploring the spatio-temporal properties of simple conducting gestures using a synchronization task. *Music Perception*, *25*, 225–239.

Luck, G., & Sloboda, J. A. (2009). Spatio-temporal cues for visually mediated synchronization. *Music Perception*, *26*, 465–473.

Lundqvist, L.-O., Carlsson, F., Hilmersson, P., & Juslin, P. N. (2009). Emotional responses to music: Experience, expression, and physiology. *Psychology of Music*, *37*, 61–90.

Luquet, G. (1913). *Les dessins d'un enfant* [The drawings of a child]. Paris: Alcan.

Lynch, M. P., & Eilers, R. E. (1992). A study of perceptual development for musical tuning. *Perception & Psychophysics*, *52*, 599–608.

Lynch, M. P., Eilers, R. E., Oller, D. K., & Urbano, R. C. (1990). Innateness, experience, and music perception. *Psychological Science*, *1*, 272–276.

Ma, W., Golinkoff, R. M., Houston, D., & Hirsh-Pasek, K. (2011). Word learning in infant- and adult-directed speech. *Language Learning and Development*, *7*, 209–225.

MacDonald, R. A. R., & Miell, D. (2000). Children's creative collaborations: The importance of friendship when working together on a musical composition. *Social Development*, *9*, 348–369.

MacDonald, R. A. R., Miell, D., & Mitchell, L. (2002). An investigation of children's musical collaborations: The effect of friendship and age. *Psychology of Music*, *30*, 148–163.

Macnamara, B. N., Hambrick, D. Z., & Oswald, F. L. (2014). Deliberate practice and performance in music, games, sports, education, and professions: A meta-analysis. *Psychological Science*, *25*, 1608–1618.

Madison, G. (2009). An auditory illusion of infinite tempo change based on multiple temporal levels. *PLOS ONE*, *4*, e8151.

Madsen, C. K. (2004). A 30-year follow-up study of actual applied music practice versus estimated practice. *Journal of Research in Music Education*, *52*, 77–88.

Magill, J. M., & Pressing, J. L. (1997). Asymmetric cognitive clock structures in West African rhythm. *Music Perception*, *15*, 189–222.

Maidhof, C., Rieger, M., Prinz, W., & Koelsch, S. (2009). Nobody is perfect: ERP effects prior to performance errors in musicians indicate fast monitoring processes. *PLOS ONE*, *4*.

Malbrán, S. (2000). The development of pulse synchrony: An exploratory study of three-year-old children. *Bulletin of the Council for Research in Music Education*, *147*, 109–115.

Malcolm, M., Lavine, A., Kenyon, G., Massie, C., & Thaut, M. (2008). Repetitive transcranial magnetic stimulation interrupts phase synchronization during rhythmic motor entrainment. *Neuroscience Letters*, *435*, 240–245.

Mangione, S., & Nieman, L. Z. (1997). Cardiac auscultatory skills of internal medicine and family practice trainees. A comparison of diagnostic proficiency. *Journal of the American Medical Association*, *278*, 717–722.

Manternach, J. N. (2012a). The effect of nonverbal conductor lip rounding and eyebrow lifting on singers' lip and eyebrow postures: A motion capture study. *International Journal of Research in Choral Singing*, *4*, 36–46.

Manternach, J. N. (2012b). The effect of varied conductor preparatory gestures on singer upper body movement. *Journal of Music Teacher Education*, *22*, 20–34.

Maquet, P., Laureys, S., Perrin, F., Ruby, P., Melchior, G., Boly, M., … Peigneux, P. (2003). Festina lente: Evidences for fast and slow learning processes and a role for sleep in human motor skill learning. *Learning & Memory*, *10*, 237–239.

Marek, C. (1975). *Toscanini*. London: Vision Press.

Margulis, E. H. (2005). A model of melodic expectation. *Music Perception*, *22*, 663–714.

Margulis, E. H. (2007). Silences in music are musical not silent: An exploratory study of context effects on the experience of musical pauses. *Music Perception*, *24*, 485–506.

Margulis, E. H. (2010). When program notes don't help: Music descriptions and enjoyment. *Psychology of Music*, *38*, 285–302.

Marin, O. S. M., & Perry, D. W. (1999). Neurological aspects of music perception and performance. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 653–724). San Diego, CA: Academic Press.

Marsh, K. (2008). *The musical playground: Global tradition and change in children's songs and games*. New York: Oxford University Press.
Marsh, K., & Young, S. (2016). Musical play. In G. E. McPherson (Ed.), *The child as musician* (pp. 462–484). Oxford: Oxford University Press.
Marshall, N. A., & Shibazaki, K. (2012). Instrument, gender and musical style associations in young children. *Psychology of Music*, *40*, 494–507.
Marshall, S. K., & Cohen, A. J. (1988). Effects of musical soundtracks on attitudes toward animated geometric figures. *Music Perception*, *6*, 95–112.
Martin, K. D., & Kim, Y. E. (1998, October). *Musical instrument identification: A pattern-recognition approach*. Presented at the 136th meeting of the Acoustical Society of America, Norfolk, VA.
Masataka, N. (2006). Preference for consonance over dissonance in hearing newborns of deaf parents and of hearing parents. *Developmental Science*, *9*, 46–50.
Maski, K. P. (2015). Sleep-dependent memory consolidation in children. *Seminars in Pediatric Neurology*, *22*, 130–134.
McAdams, S. (2013). Musical timbre perception. In D. Deutsch (Ed.), *The psychology of music* (3rd ed., pp. 35–68). Amsterdam: Academic Press.
McAuley, J. D., Jones, M. R., S., H., Johnston, H. M., & Miller, N. S. (2006). The time of our lives: Lifespan development of timing and event tracking. *Journal of Experimental Psychology: General*, *135*, 348–367.
McDermott, J., & Hauser, M. D. (2004). Are consonant intervals music to their ears? Spontaneous acoustic preferences in a nonhuman primate. *Cognition*, *94*, B11–B21.
McDermott, J., & Hauser, M. (2005). The origins of music: Innateness, uniqueness, and evolution. *Music Perception*, *23*, 29–60.
McDermott, J., & Hauser, M. D. (2007). Nonhuman primates prefer slow tempos but dislike music overall. *Cognition*, *104*, 654–668.
McDermott, J. H., Lehr, A. J., & Oxenham, A. J. (2010). Individual differences reveal the basis of consonance. *Current Biology*, *20*, 1035–1041.
McDermott. J. H., Schultz, A. F., Undurraga, E. A., & Godoy, R. A. (2016). Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature*, *535*, 547–555.
McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746–748.
McKeage, K. M. (2004). Gender and participation in high school and college instrumental jazz ensembles. *Journal of Research in Music Education*, *52*, 343–356.
McKelvie, P., & Low, J. (2002). Listening to Mozart does not improve children's spatial ability: Final curtains for the Mozart effect. *British Journal of Developmental Psychology*, *20*, 241–258.
McLachlan, N., Marco, D., Light, M., & Wilson, S. (2013). Consonance and pitch. *Journal of Experimental Psychology: General*, *142*, 1142–1158.
McLeod, P., Plunkett, K., & Rolls, E. T. (1998). *Introduction to connectionist modeling of cognitive processes*. Oxford: Oxford University Press.
McMullen, E., & Saffran, J. R. (2004). Music and language: A developmental comparison. *Music Perception*, *21*, 289–311.
McNamara, T. P. (2005). *Semantic priming: Perspectives from memory and word recognition*. London: Routledge.
McPherson, G. E. (1994). Factors and abilities influencing sightreading skill in music. *Journal of Research in Music Education*, *42*, 217–231.
McPherson, G. E. (2005). From child to musician: Skill development during the beginning stages of learning an instrument. *Psychology of Music*, *33*, 5–35.
McPherson, G. E. (2009). The role of parents in children's musical development. *Psychology of Music*, *37*, 91–110.
McPherson, G. E., & McCormick, J. (1999). Motivational and self-regulated learning components of musical practice. *Bulletin of the Council for Research in Music Education*, *141*, 98–102.
McPherson, G. E., & Renwick, J. M. (2001). A longitudinal study of self-regulation in children's musical practice. *Music Education Research*, *3*, 169–186.

McPherson, G. E., & Zimmerman, B. J. (2011). Self-regulation of musical learning: A social cognitive perspective on developing performance skills. In R. Colwell & P. Webster (Eds.), *MENC handbook of research on music learning, Volume 2: Applications* (pp. 130–175). New York: Oxford University Press.

McPherson, M. J., Lopez-Gonzalez, M., Rankin, S. K., & Limb, C. J. (2014). The role of emotion in musical improvisation: An analysis of structural features. *PLOS ONE*, *9*, e105144.

Mehr, S. A., Song, L. A., & Spelke, E. S. (2016). For 5-month-old infants, melodies are social. *Psychological Science*, *27*, 486–501.

Meyer, J. (2015). Reflections on the spatial-sound imagination of great composers. *Psychomusicology: Music, Mind, and Brain*, *25*, 187–194.

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

Meyer, R., Palmer, C., & Mazo, M. (1998). Affective and coherence responses to Russian laments. *Music Perception*, *16*, 135–150.

Miklaszewski, K. (1989). A case study of a pianist preparing a musical performance. *Psychology of Music*, *17*, 95–109.

Millar, B. (2008). Selective hearing: Gender bias in the music preferences of young adults. *Psychology of Music*, *36*, 429–445.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97.

Milliman, R. E. (1982). Using background music to affect the behavior of supermarket shoppers. *Journal of Marketing*, *46*, 86–91.

Mishra, J. (2002). A qualitative analysis of strategies employed in efficient and inefficient memorization. *Bulletin of the Council for Research in Music Education*, *152*, 74–86.

Mitani, C., Nakata, T., Treub, S. E., Kanda, Y., Kumagami, H., Takasaki, K., … Takahashi, H. (2007). Music recognition, music listening, and word recognition by deaf children with cochlear implants. *Ear and Hearing*, *28*, 29S–33S.

Mithen, S. (2006). *The singing Neanderthals: The origins of music, language, mind and body*. Cambridge, MA: Harvard University Press.

Mitterschiffthaler, M. T., Fu, C. H. Y., Dalton, J. A., Andrew, C. M., & Williams S. C. R. (2007). A functional MRI study of happy and sad affective states induced by classical music. *Human Brain Mapping*, *28*, 1150–1162.

Miyazaki, K. (2004). Recognition of transposed melodies by absolute-pitch possessors. *Japanese Psychological Research*, *46*, 270–282.

Miyazaki, K., Makomaska, S., & Rakowski, A. (2012). Prevalence of absolute pitch: A comparison between Japanese and Polish music students. *Journal of the Acoustical Society of America*, *132*, 3484–3493.

Moelants, D. (2002). Preferred tempo reconsidered. In C. Stevens, D. Burnham, G. McPherson, E. Schubert, & J. Renwick (Eds.), *Proceedings of the 7th International Conference on Music Perception and Cognition* (pp. 580–583). Adelaide: Causal Productions.

Moog, H. (1976a). The development of musical experience in children of pre-school age. *Psychology of Music*, *4*, 38–45.

Moog, H. (1976b). *The musical experience of the pre-school child* (C. Clarke, Trans.). London: Schott.

Moore, B. J. C. (2012). *An introduction to the psychology of hearing* (6th ed.). San Diego, CA: Academic Press.

Moorhead, G. E., & Pond, D. (1941). *Music of young children: I. Chant*. Santa Barbara, CA: Pillsbury Foundation for Advancement of Music Education.

Moorhead, G. E., & Pond, D. (1942). *Music of young children: II. General observations*. Santa Barbara, CA: Pillsbury Foundation for Advancement of Music Education.

Moorhead, G. E., Sandvik, F., & Wight, D. (1951). *Music of young children: IV. Free use of instruments for musical growth*. Santa Barbara, CA: Pillsbury Foundation for Advancement of Music Education.

Morrison, S. J., & Fyk, J. (2002). Intonation. In R. Parncutt & G. E. McPherson (Eds.), *The science and psychology of music performance* (pp. 183–197). Oxford: Oxford University Press.

Morrison, S. J., Montemayor, M., & Wiltshire, E. S. (2004). The effect of a recorded model on band students' performance self-evaluations, achievement, and attitude. *Journal of Research in Music Education*, *52*, 116–129.
Morrison, S. J., Price, H. E., Geiger, C. G., & Cornacchio, R. A. (2009). The effect of conductor expressivity on ensemble performance evaluation. *Journal of Research in Music Education*, *57*, 37–49.
Morrison, S. J., Price, H. E., Smedley, E. M., & Meals, C. D. (2014). Conductor gestures influence evaluations of ensemble performance. *Frontiers in Psychology*, *5*, 806.
Morrison, S., & Selvey, J. (2014). The effect of conductor expressivity on choral ensemble evaluation. *Bulletin of the Council for Research in Music Education*, *99*, 7–18.
Morrongiello, B. A., & Rocca, P. T. (1987). Infants' localization of sounds in the horizontal plane: Effects of auditory and visual cues. *Child Development*, *58*, 918–927.
Morton, J. B., & Trehub, S. E. (2007). Children's judgments of emotion in song. *Psychology of Music*, *35*, 629–639.
Mosing, M. A., Pedersen, N. L., Madison, G., & Ullén, F. (2014). Genetic pleiotropy explains associations between musical auditory discrimination and intelligence. *PLOS ONE*, *9*, e113874.
Mualem, O., & Klein, P. S. (2013). The communicative characteristics of musical interactions compared with play interactions between mothers and their one-year-old infants. *Early Child Development and Care*, *183*, 899–915.
Muir, D., & Field, J. (1979). Newborn infants orient to sounds. *Child Development*, *50*, 431–436.
Müllensiefen, D., & Halpern, A. R. (2014). Higher-order and lower-order features through machine learning. *Music Perception*, *31*, 418–435.
Mürbe, D., Pabst, F., Hofmann, G., & Sundberg, J. (2003). Effects of a professional solo singer education on auditory and kinesthetic feedback – a longitudinal study of singers' pitch control. *Journal of Voice*, *18*, 236–241.
Murnighan, J. K., & Conlon, D. E. (1991). The dynamics of intense work groups: A study of British string quartets. *Administrative Science Quarterly*, *36*, 165–186.
Mutschler, I., Schultze-Bonhage, A., Glauche, V., Demandt, E., Speck, O., & Ball, T. (2007). A rapid sound–action association effect in human insular cortex. *PLOS ONE*, *2*, e259.
Nager, W., Kohlmetz, C., Altenmüller, E., Rodriguez-Fornells, A., & Münte, T. F. (2003). The fate of sounds in conductors' brains: An ERP study. *Cognitive Brain Research*, *17*, 83–93.
Nakata, T., Trehub, S. E., Mitani, C., & Kanda, Y. (2006). Pitch and timing in the songs of deaf children with cochlear implants. *Music Perception*, *24*, 147–154.
Nantais, K. M., & Schellenberg, E. G. (1999). The Mozart effect: An artifact of preference. *Psychological Science*, *10*, 370–373.
Narmour, E. (1990). *The analysis and cognition of basic melodic structures: The Implication–Realization model*. Chicago, IL: University of Chicago Press.
Narmour, E. (1991). The top-down and bottom-up systems of musical implication: Building on Meyer's theory of emotional syntax. *Music Perception*, *9*, 1–26.
Narmour, E. (1992). *The analysis and cognition of melodic complexity: The Implication–Realization model*. Chicago, IL: University of Chicago Press.
Nettl, B. (2001). Studying music of the world's cultures. In B. Nettl, C. Capwell, I. K. F. Wong, & T. Turino (Eds.), *Excursions in world music* (pp. 1–18). Upper Saddle River, NJ: Prentice Hall.
Nicholson, D. R., Cody, M. W., & Beck, J. G. (2015). Anxiety in musicians: On and off stage. *Psychology of Music*, *43*, 438–449.
Nilsson, B., & Folkestad, G. (2005). Children's practice of computer-based composition. *Music Education Research*, *7*, 21–37.
Nöcker-Ribaupierre, M. (2004). *Music therapy for premature and newborn infants*. Gilsum, NH: Barcelona Publishers.
Nombela, C., Hughes, L. E., Owen, A. M., & Grahn, J. A. (2013). Into the groove: Can rhythm influence Parkinson's disease? *Neuroscience & Biobehavioral Reviews*, *37*, 2564–2570.
Norgaard, M. (2014). How jazz musicians improvise: The central role of auditory and motor patterns. *Music Perception*, *31*, 271–287.

Norgaard, M., Emerson, S. N., Dawn, K., & Fidlon, J. D. (2016). Creating under pressure: Effects of divided attention on the improvised output of skilled jazz pianists. *Music Perception*, *33*, 561–570.

Norman-Haignere, S., Kanwisher, N. G., & McDermott, J. H. (2015). Distinct cortical pathways for music and speech revealed by hypothesis-free voxel decomposition. *Neuron*, *88*, 1281–1296.

North, A. C., Colley, A. M., & Hargreaves, D. J. (2003). Adolescents' perceptions of the music of male and female composers. *Psychology of Music*, *31*, 139–154.

North, A. C., & Hargreaves, D. J. (1998). The effect of music on atmosphere and purchase intentions in a cafeteria. *Journal of Applied Social Psychology*, *28*, 2254–2273.

North, A. C., & Hargreaves, D. J. (Eds.). (2008). *The social and applied psychology of music*. Oxford: Oxford University Press.

North, A. C, Hargreaves, D. J., & Krause, A. E. (2016). Music and consumer behavior. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford handbook of music psychology* (2nd ed., pp. 789–801). Oxford: Oxford University Press.

North, A. C., Hargreaves, D. J., & McKendrick, J. (1997). In-store music affects product choice. *Nature*, *390*, 132.

Nusseck, M., & Wanderley, M. M. (2009). Music and motion: How music-related ancillary body movements contribute to the experience of music. *Music Perception*, *26*, 335–353.

Occupational Safety and Health Administration (OSHA) (1983). *Occupational noise exposure; hearing conservation amendment: Final rule* (Fed. Reg. 48:9738-9785). Washington, DC: US Department of Labor.

O'Flynn, J. (2006). Vernacular music-making and education. *International Journal of Music Education*, *24*, 140–147.

Olsho, L. W., Koch, E. G., Halpin, C. F. (1987). Level and age effects in infant frequency discrimination. *Journal of the Acoustical Society of America*, *82*, 454–464.

O'Neill, C. T., Trainor, L. J., & Trehub, S. E. (2001). Infants' responsiveness to fathers' singing. *Music Perception*, *18*, 409–425.

O'Neill, S. A. (1997a). Gender and music. In D. J. Hargreaves & A. C. North (Eds.), *The social psychology of music* (pp. 46–66). Oxford: Oxford University Press.

O'Neill, S. A. (1997b). The role of practice in children's early musical performance achievement. In H. Jørgensen & A. C. Lehmann (Eds.), *Does practice make perfect? Current theory and research on instrumental music practice* (pp. 53–70). Oslo: Norges Musikkhøgskole.

O'Neill, S. A., & Boulton, M. J. (1996). Boys' and girls' preferences for musical instruments. *Psychology of Music*, *24*, 171–183.

Osterhout, L., & Holcomb, P. J. (1992). Event-related potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*, 785–806.

Ostri, B., Eller, N., Dahlin, E., & Skylv, G. (1989). Hearing impairment in orchestral musicians. *Scandinavian Audiology*, *18*, 243–249.

Overy, K., Norton, A., Cronin, K., Winner, E., & Schlaug, G. (2005). Examining rhythm and melody processing in young children using fMRI. In G. Avanzini, S. Koelsch, L. Lopez, & M. Majno (Eds.), *The neurosciences and music II: From perception to performance* (pp. 210–218). New York: New York Academy of Sciences.

Oxendine, J. B. (1984). *Psychology of motor learning* (2nd ed.). New York: Appleton-Century-Crofts.

Ozon, F. (Writer/Director), Delbosc, O., & Missonier, M. (Producers). (2003). *Swimming pool* [Motion picture]. Paris: Fidélité Productions. (Available from Focus Features, 100 Universal City Plaza, University City, CA 91608)

Palmer, C. (1989). Mapping musical thought to musical performance. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 31–346.

Palmer, C. (1997). Music performance. *Annual Review of Psychology*, *48*, 115–138.

Palmer, C., & Krumhansl, C. L. (1987). Independent temporal and pitch structures in determination of musical phrases. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 116–126.

Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 728–741.
Palmer, C., & Pfordresher, P. Q. (2003). Incremental planning in sequence production. *Psychological Review*, *110*, 683–712.
Palmer, C., & van de Sande, C. (1993). Units of knowledge in music performance. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *19*, 457–470.
Palmer, C., & van de Sande, C. (1995). Range of planning in music performance. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 947–962.
Paney, A. S., & Kay, A. C. (2015). Developing singing in third-grade music classrooms: The effect of a concurrent-feedback computer game on pitch-matching skills. *Update: Applications of Research in Music Education*, *34*, 42–49.
Panneton, R. K. (1985). *Prenatal auditory experience with melodies: Effects on postnatal auditory preferences in human newborns*. Unpublished doctoral dissertation, University of North Carolina at Greensboro, North California.
Pantev, C., Engelien, A., Candia, V., & Elbert, T. (2003). Representational cortex in musicians. In I. Peretz & R. Zatorre (Eds.), *The cognitive neuroscience of music* (pp. 382–395). New York: Oxford University Press.
Papageorgi, I., Creech, A., & Welch, G. (2013). Perceived performance anxiety in advanced musicians specializing in different musical genres. *Psychology of Music*, *41*, 18–41.
Papoušek, M., Papoušek, H., & Symmes, D. (1991). The meanings of melodies in motherese in tone and stress languages. *Infant Behavior & Development*, *14*, 415–440.
Park, M.-Y. (2003). Assessment of potential noise-induced hearing loss with commercial 'karaoke' noise. *International Journal of Industrial Ergonomics*, *31*, 375–385.
Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, *11*, 409–464.
Parsons, L. M., Sergent, J., Hodges, D. A., & Fox, P. T. (2005). The brain basis of piano performance. *Neuropsychologia*, *43*, 119–215.
Partanen, E., Kujala, T., Tervaniemi, M., & Huotilainen, M. (2013). Prenatal music exposure induces long-term neural effects. *PLOS ONE*, *8*, e78946.
Pascual-Leone, A. (2003). The brain that makes music and is changed by it. In I. Peretz & R. Zatorre (Eds.), *The cognitive neuroscience of music* (pp. 396–409). New York: Oxford University Press.
Pascual-Leone, A., Nguyet, D., Cohen, L. G., Brasil-Neto, J. P., Cammarota, A., & Hallett, M. (1995). Modulation of muscle responses evoked by transcranial magnetic stimulation during the acquisition of new fine motor skills. *Journal of Neurophysiology*, *74*, 1037–1045.
Patel, A. D. (2008). *Music, language, and the brain*. New York: Oxford.
Patel, A. D. (2014). Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. *Hearing Research*, *308*, 98–108.
Patel, A. D. (2015). *Music and the brain*. Chantilly, VA: The Great Courses. [Set of eighteen 30-minute lectures available in video or audio format]
Patel, A. D., & Daniele, J. R. (2003). An empirical comparison of rhythm in language and music. *Cognition*, *87*, B35–B45.
Patel, A. D., & Demorest, S. D. (2013). Comparative music cognition: Cross-species and cross-cultural studies. In D. Deutsch (Ed.), *The psychology of music* (3rd ed., pp. 647–682). Amsterdam: Academic Press.
Patel, A. D., Gibson, E., Ratner, J., Besson, M., Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study *Journal of Cognitive Neuroscience*, *10*, 717–733.
Patel, A. D., & Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: The action simulation for auditory prediction (ASAP) hypothesis. *Frontiers in Systems Neuroscience*, *8*, 57.
Patel, A. D., Iversen, J. R., Bregman, M. R., & Schultz, I. (2009). Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Current Biology*, *19*, 1–4.

Pätynen, J., Tervo, S., Robinson, P. W., & Lokki, T. (2014). Concert halls with strong lateral reflections enhance musical dynamics. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 4409–4414.

Peng J. H., Tao, Z. Z., & Huang, Z. W. (2007). Risk of damage to hearing from personal listening devices in young adults. *Journal of Otolaryngology*, *36*, 181–185.

Peretz, I. (1990). Processing of local and global musical information by unilateral brain-damaged patients. *Brain*, *113*, 1185–1205.

Peretz, I. (1993). Auditory atonalia for melodies. *Cognitive Neuropsychology*, *10*, 21–56.

Peretz, I. (2003). Brain specialization for music: new evidence from congenital amusia. In I. Peretz & R. Zatorre (Eds.), *The cognitive neuroscience of music* (pp. 192–203). New York: Oxford University Press.

Peretz, I. (2012). Music, language, and modularity in action. In P. Rebruschat, M. Rohrmeir, J. A. Hawkins, & I. Cross (Eds.), *Language and music as cognitive systems* (pp. 254–268). Oxford: Oxford University Press.

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, *6*, 688–691.

Peretz, I., Vuvan, D. T., Lagrois, M.-E., & Armony, J. L. (2015). Neural overlap in processing music and speech. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, *370*, 20140090.

Peretz, I., & Zatorre, R. J. (2005). Brain organization for music processing. *Annual Review of Psychology*, *56*, 89–114.

Perlman, M., & Krumhansl, C. L. (1996). An experimental study of internal interval standards in Javanese and Western Musicians. *Music Perception*, *14*, 95–116.

Petacchi, A., Laird, A. R., Fox, P. T., & Bower, J. M. (2005). Cerebellum and auditory function: An ALE meta-analysis of functional neuroimaging studies. *Human Brain Mapping*, *25*, 105–117.

Peterson, B., Mortensen, M. V., Hansen, M., & Vuust, P. (2012). Singing in the key of life – a pilot study on effects of musical ear training after cochlear implantation. *Psychomusicology: Music, Mind, & Brain*, *22*, 134–151.

Pfordresher, P. Q. (2003). Auditory feedback in music performance: Evidence for a dissociation of sequencing and timing. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 949–964.

Pfordresher, P. Q. (2006). Coordination of perception and action in music performance. *Advances in Cognitive Psychology*, *2*, 183–198.

Pfordresher, P. Q., & Beasley, R. T. E. (2014). Making and monitoring errors based on altered auditory feedback. *Frontiers in Psychology*, *5*, 914.

Pfordresher, P. Q., & Brown, S. (2007). Poor-pitch singing in the absence of 'tone deafness.' *Music Perception*, *25*, 95–115.

Pfordresher, P. Q., & Larrouy-Maestri, P. (2015). On drawing a line through the spectrogram: How do we understand deficits of vocal pitch imitation? *Frontiers in Human Neuroscience*, *9*, 271.

Pfordresher, P. Q., & Mantell, J. T. (2012). Effects of altered auditory feedback across effector systems: Production of melodies by keyboard and singing. *Acta Psychologica*, *139*, 166–177.

Pfordresher, P. Q., Palmer, C., & Jungers, M. (2007). Speed, accuracy and serial order in sequence production. *Cognitive Science*, *31*, 63–98.

Phelps, A. L. (2010). *Beyond auditions: Gender discrimination in America's top orchestras* (Doctoral dissertation). Retrieved from Iowa Research Online.

Phillips, S. L., Henrich, V. C., & Mace, S. T. (2010). Prevalence of noise-induced hearing loss in student musicians. *International Journal of Audiology*, *49*, 309–316.

Phillips, S. L., & Mace, S. (2008). Sound level measurements in music practice rooms. *Music Performance Research*, *2*, 36–47.

Phillips-Silver, J. (2014). So you think you can't dance? (The mysterious case of the guy with two left feet). *Frontiers for Young Minds*, *2*, 11.

Phillips-Silver, J., Toiviainen, P., Gosselin, N., Piche, O., Nozaradan, S., Palmer, C., … Peretz, I. (2011). Born to dance but beat deaf: A new form of congenital amusia. *Neuropsychologia*, *49*, 961–969.

Phillips-Silver, J., Toiviainen, P., Gosselin, N., Turgeon, C., Lepore, F., & Peretz, I. (2015). Cochlear implant users move in time to the beat of drum music. *Hearing Research*, *321*, 25–34.

Phillips-Silver, J., & Trainor, L. J. (2005, June 3). Feeling the beat: Movement influences infant rhythm perception. *Science*, *308*, 1430.

Phillips-Silver, J., & Trainor, L. J. (2007). Hearing what the body feels: Auditory encoding of rhythmic movement. *Cognition*, *105*, 533–546.

Piaget, J. (1936). *Origins of intelligence in the child.* London: Routledge & Kegan Paul.

Pickering, S., & Repacholi, B. (2001). Modifying children's gender-typed musical instrument preferences: The effects of gender and age. *Sex Roles*, *45*, 623–643.

Pickles, J. O. (2008). *An introduction to the physiology of hearing* (3rd ed.). Bingley: Emerald Books.

Pierce, J. R. (1992). *The science of musical sound*. New York: Freeman.

Pike, K. (1945). *The intonation of American English*. Ann Arbor, MI: University of Michigan Press.

Pinker, S. (1997). *How the mind works*. New York: W. W. Norton.

Pitts, S. E., Dobson, M. C., Gee, K., & Spencer, C. P. (2013). Views of an audience: Understanding the orchestral concert experience from player and listener perspectives. *Participations: Journal of Audience & Reception Studies*, *10*, 65–95.

Plack, C. J. (2014). *The sense of hearing* (2nd ed.). New York: Routledge/Taylor & Francis.

Plantinga, J., & Trainor, L. J. (2005). Memory for melody: Infants use a relative pitch code. *Cognition*, *98*, 1–11.

Plantinga, J., & Trehub, S. E. (2014). Revisiting the innate preference for consonance. *Journal of Experimental Psychology: Human Perception and Performance*, *40*, 40–49.

Platz, F., Kopiez, R., Lehmann, A. C., & Wolf, A. (2014). The influence of deliberate practice on musical achievement: A meta-analysis. *Frontiers in Psychology*, *5*, 646.

Plomp, R., & Levelt, W. J. M. (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America*, *38*, 548–560.

Portnuff, C. D. (2016). Reducing the risk of music-induced hearing loss from overuse of portable listening devices: Understanding the problems and establishing strategies for improving awareness in adolescents. *Adolescent Health, Medicine and Therapeutics*, *7*, 27–35.

Povel, D. J. (1981). Internal representation of simple temporal patterns. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 3–18.

Povel, D. J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception*, *2*, 411–441.

Pratt, C. C. (1931). *The meaning of music: A study in psychological aesthetics* (1st ed.). New York: McGraw-Hill.

Prensky, M. (2001). Digital natives, digital immigrants. *On the Horizon*, *9*, 1–6.

Prim, F. M. (1989). The importance of girls' singing games in music and motor education. *Canadian Journal of Research in Music Education*, *32*, 115–123.

Prince, J. B., Schmuckler, M. A., & Thompson, W. F. (2009). The effect of task and pitch structure on pitch–time interactions in music. *Memory & Cognition*, *37*, 368–381

Profita, J., & Bidder, T. G. (1988). Perfect pitch. *American Journal of Medical Genetics*, *29*, 763–771.

Profyt, L., & Whissell, C. (1991). Children's understanding of facial expression of emotion: I. Voluntary creation of emotion-faces. *Perceptual and Motor Skills*, *73*, 199–202.

Pujol, R., Laville-Rebillard, M., & Lenoir, M. (1998). Development of sensory and neural structures in the mammalian cochlea. In E. W. Rubel, A. N. Popper, & R. R. Fay (Eds.), *Development of the auditory system* (pp. 146–193). New York: Springer-Verlag.

Quarrier, N. F. (1993). Forward head posture in vocal performance. *Medical Problems of Performing Artists*, *8*, 29–32.

Querleu, D., Lefebre, C., Renard, X., Titran, M., Morillion, M., & Crépin, G. (1984). Réactivité du nouveau-né de moins de deux heures de vie a la voix maternelle [Responsivity of the newborn to the maternal voice at less than two hours old]. *Journal de Gynécologie, Obstétrique et Biologie de la Reproduction*, *13*, 125–134.

Querleu, D., Renard, X., Boutteville, C., & Crépin, G. (1989). Hearing by the human fetus? *Seminars in Perinatology*, *13*, 409–420.

Querleu, D., Renard, X., Versyp, F., Paris-Delrue, L., & Crépin, G. (1988). Fetal hearing. *European Journal of Obstetrics and Gynecology and Reproductive Biology*, *29*, 191–212.

Quinn, I. (1999). The combinatorial model of pitch contour. *Music Perception*, *16*, 439–456.

Quinto, L. R., Thompson, W. F., Kroos, C., & Palmer, C. (2014). Singing emotionally: A study of pre-production, production, and post-production facial expressions. *Frontiers in Psychology*, *5*, 262.

Radocy, R. E., & Boyle, J. D. (2003). *Psychological foundations of musical behavior*. Springfield, IL: Thomas.

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences*, *98*, 676–682.

Rainbow, E. L. (1981). A final report on a three-year investigation of the rhythmic abilities of preschool-aged children. *Bulletin of the Council for Research in Music Education*, *66–67*, 69–73.

Raman, R., & Dowling, W. J. (2016). Real-time probing of modulations in South Indian classical (Carnātic) music by Indian and Western musicians. *Music Perception*, *33*, 367–391.

Ramella, M., Fronte, F., & Converti, R. M. (2014). Postural disorders in conservatory students: The Diesis project. *Medical Problems of Performing Artists*, *29*, 19–22.

Rand, K., & Lahav, A. (2014). Maternal sounds elicit sustained heart rate response in preterm newborns in the first month of life. *Early Human Development*, *90*, 679–683.

Randall, W. M., & Rickard, N. S. (2017). Personal music listening: A model of emotional outcomes developed through Mobile Experience Sampling. *Music Perception*, *34*, 501–514.

Rasch, R., & Plomp, R. (1999). The perception of musical tones. In D. Deutsch (Ed.), *The psychology of music* (pp. 89–112). San Diego, CA: Academic Press.

Rashotte, M. A., & Wedell, D. H. (2014). Testing the absolute-tempo hypothesis: Context effects for familiar and unfamiliar songs. *Memory & Cognition*, *42*, 1302–1314.

Rauschecker, J. P., & Tian, B. (2000). Mechanisms and streams for processing of 'what' and 'where' in auditory cortex. *Proceedings of the National Academy of Sciences*, *97*, 11800–11806.

Rauscher, F. H., & Hinton, S. C. (2006). The Mozart effect: Music listening is not music instruction. *Educational Psychologist*, *41*, 233–238.

Rauscher, F. H., Shaw, G. L., & Ky, K. N. (1993). Music and spatial task performance. *Nature*, *365*, 6447.

Rauscher, F. H., Shaw, G. L., & Ky, K. N. (1995). Listening to Mozart enhances spatial-temporal reasoning: Towards a neurophysiological basis. *Neuroscience Letters*, *185*, 44–47.

Reber, A. S. (1985). *The Penguin dictionary of psychology*. London: Penguin Books.

Reichardt, W., Alim, A., & Schmidt, W. (1974). Abhängigkeit der Grenzen zwischen brauchbarer und unbrauchbarer Durchsichtigkeit von der Art des Musikmotives, der Nachhallzeit und der Nachhalleinsatzzeit. *Applied Acoustics*, *7*, 243–264.

Reigado, J., Rocha, A., & Rodrigues, H. (2011). Vocalizations of infants (9–11 months old) in response to musical and linguistic stimuli. *International Journal of Music Education*, *29*, 241–256.

Repp, B. H. (1991). Diversity and commonality in music performance: An analysis of timing microstructure in Schumann's 'Träumerei.' *Journal of the Acoustical Society of America*, *92*, 2546–2568.

Repp, B. H. (1994a). On determining the basic tempo of an expressive music performance. *Psychology of Music*, *22*, 157–167.

Repp, B. H. (1994b). Relational invariance of expressive microstructure across global tempo changes in music performance: An exploratory study. *Psychological Research*, *56*, 269–284.

Repp, B. H. (1996). The art of inaccuracy: Why pianists' errors are difficult to hear. *Music Perception*, *14*, 161–183.

Repp, B. H. (1999). Effects of auditory feedback deprivation on expressive piano performance. *Music Perception*, *16*, 409–438.

Repp, B. H. (2000). Pattern typicality and dimensional interactions in pianists' imitation of expressive timing and dynamics. *Music Perception*, *18*, 173–212.

Repp, B. H. (2011). Comments on ‘Tactus ≠ Tempo: Some dissociations between attentional focus, motor behavior, and tempo judgment’ by Justin London. *Empirical Musicology Review*, *6*, 56–61.

Repp, B. H., London, J., & Keller, P. E. (2012). Distortions in reproduction of two-interval rhythms: When the ‘attractor ratio’ is not exactly 1:2. *Music Perception*, *30*, 205–223.

Révész, G. (1954). *Introduction to the psychology of music* (C. I. C. de Courcy, Trans.). Norman, OK: University of Oklahoma.

Rickert, D. L. L., Barrett, M. S., & Ackermann, B. J. (2014). Injury and the orchestral environment part II: Organisational culture, behavioural norms, and attitudes to injury. *Medical Problems of Performing Artists*, *29*, 94–101.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, *27*, 169–192.

Roballey, T. C., McGreevy, C., Rongo, R. R., Schwantes, M. L., Steger, P. J., Wininger, M. A., & Gardner, E. B. (1985). The effect of music on eating behavior. *Bulletin of the Psychonomic Society*, *23*, 221–222.

Roberts, S., Fyfield, R., Baibazarova, E., van Goozen, S., Culling, J. F., & Hay, D. F. (2013). Parental speech at 6 months predicts joint attention at 12 months. *Infancy*, *18*, E1–E15.

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. Oxford: Oxford University Press.

Rosenkranz, K., Williamon, A., Butler, K., Cordivari, C., Lees, A. J., & Rothwell, J. C. (2005). Pathophysiological differences between musician’s dystonia and writer’s cramp. *Brain*, *128*, 918–931.

Rosenthal, R. K. (1984). Relative effects of guided model, model only, guide only, and practice only treatments on the accuracy of advanced instrumentalists’ musical performance. *Journal of Research in Music Education*, *3*, 265–273.

Rosner, B. S., & Meyer, L. B. (1982). Melodic processes and the perception of music. In D. Deutsch (Ed.), *The psychology of music* (pp. 317–343). San Diego, CA: Academic Press.

Ross, S. L. (1985). The effectiveness of mental practice in improving the performance of college trombonists. *Journal of Research in Music Education*, *33*, 221–230.

Royal, I., Vuvan, D. T., Zendel, B. R., Robitaille, N., Schönwiesner, M., & Peretz, I. (2016). Activation in the right inferior parietal lobule reflects the representation of musical structure beyond simple pitch discrimination. *PLOS ONE*, *11*, e0155291.

Rubin-Rabson, G. (1940a). Studies in the psychology of memorizing music: II. A comparison of massed and distributed learning. *Journal of Educational Psychology*, *31*, 270–284.

Rubin-Rabson, G. (1941a). Studies in the psychology of memorizing music: V. A comparison of pre-study periods of varied length. *Journal of Educational Psychology*, *32*, 101–112.

Rubin-Rabson, G. (1941b). Studies in the psychology of memorizing music: VI. A comparison of two forms of mental rehearsal and keyboard overlearning. *Journal of Educational Psychology*, *32*, 593–602.

Ruiz, M. H., Jabusch, H. C., & Altenmüller, E. (2009). Detecting wrong notes in advance: Neuronal correlates of error monitoring in pianists. *Cerebral Cortex*, *19*, 2625–2639.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*, 1161–1178.

Russo, F. A., Windell, D. L., & Cuddy, L. L. (2003). Learning the special note: Evidence for a critical period for absolute pitch acquisition. *Music Perception*, *21*, 51–62.

Ruthsatz, J., Detterman, D., Griscom, W. S., & Cirullo, B. A. (2008). Becoming an expert in the musical domain: It takes more than just practice. *Intelligence*, *36*, 330–338.

Ruthsatz, J., Ruthsatz, K., & Stephens, K. R. (2014). Putting practice into perspective: Child prodigies as evidence of innate talent. *Intelligence*, *45*, 60–65.

Rutkowski, J. (1997). The nature of children’s singing voices. In B. A. Roberts (Ed.), *The phenomenon of singing* (pp. 201–209). St. John’s, Newfoundland: Memorial University Press.

Rutkowski, J. (2015). The relationship between children’s use of singing voice and singing accuracy. *Music Perception*, *32*, 283–292.

Saarikallio, S. (2011). Music as emotional self-regulation throughout adulthood. *Psychology of Music*, *39*, 307–327.

Sabine, W. C. (1922). *Collected papers on acoustics*. Cambridge, MA: Harvard University Press.

Sachs, M. E., Damasio, A., & Habibi, A. (2015). The pleasures of sad music: A systematic review. *Frontiers in Human Neuroscience*, *9*, 404.

Sack, K. (1998, January 15). Georgia's governor seeks musical start for babies. *The New York Times*, p. A12.

Sacks, O. (2007). *Musicophilia: Tales of music and the brain*. New York: Knopf.

Sadler, M. E., & Miller, C. J. (2010). Performance anxiety: A longitudinal study of the roles of personality and experience in musicians. *Social Psychological and Personality Science*, *1*, 280–287.

Saffle, M. (1994). *Liszt in Germany 1840–1845*. New York: Pendragon Press.

Saffran, J. R. (2003). Absolute pitch in infancy and adulthood: The role of tonal structure. *Developmental Science*, *6*, 37–49.

Saffran, J. R., & Griepentrog, G. J. (2001). Absolute pitch in infant auditory learning: Evidence for developmental reorganization. *Developmental Psychology*, *37*, 74–85.

Saffran, J. R., Loman, M. M., & Robertson, R. R. W. (2000). Infant memory for musical experiences. *Cognition*, *77*, B15–B23.

Saint-Georges, C., Chetouani, M., Cassel, R., Apicella, F., Mahdhaoui, A., Muratori, F., … Cohen, D. (2013). Motherese in interaction: At the cross-road of emotion and cognition? (A systematic review). *PLOS ONE*, *8*, e78103.

Sakai, K., Hikosaka, O., Miyauchi, S., Takino, R., Tamada, T., Iwata, N. K., & Nielsen, M. (1999). Neural representation of rhythm depends on its interval ratio. *Journal of Neuroscience*, *19*, 10074–10081.

Saldanha, F. I., & Corso, J. F. (1964). Timbre cues and the identification of musical instruments. *Journal of the Acoustical Society of America*, *36*, 2021–2026.

Salimpoor, V. N., Benovoy, M., Larcher, K., Dagher, A., & Zatorre, R. J. (2011). Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nature Neuroscience*, *14*, 257–262.

Salk, L. (1961). Mothers' heartbeat as an imprinting stimulus. *Transactions of the New York Academy of Sciences*, *2*, 753–763.

Samson, S., & Ehrlé, N. (2003). Cerebral substrates for musical temporal processes. In I. Peretz & R. Zatorre (Eds.), *The cognitive neuroscience of music* (pp. 204–216). New York: Oxford University Press.

Santesso, D. L., Schmidt, L. A., & Trainor, L. J. (2007). Frontal brain electrical activity (EEG) and heart rate in response to affective infant-directed (ID) speech in 9-month-old infants. *Brain and Cognition*, *65*, 14–21.

Sataloff, R. T. (1991). Hearing loss in musicians. *American Journal of Otology*, *12*, 122–127.

Satoh, M., Takeda, K., & Kuzuhara, S. (2007). A case of auditory agnosia with impairment of perception and expression of music: Cognitive processing of tonality. *European Neurology*, *58*, 70–77.

Satt, B. J. (1984). *An investigation into the acoustical induction of intra-uterine learning*. Unpublished doctoral dissertation. Los Angeles, CA: California School of Professional Psychologists.

Schellenberg, E. G. (1996). Expectancy in melody: Tests of the Implication–Realization model. *Cognition*, *58*, 75–125.

Schellenberg, E. G. (1997). Simplifying the Implication–Realization model of melodic expectancy. *Music Perception*, *14*, 295–318.

Schellenberg, E. G. (2006). Long-term positive associations between music lessons and IQ. *Journal of Educational Psychology*, *98*, 457–468.

Schellenberg, E. G. (2012). Cognitive performance after music listening: A review of the Mozart effect. In R. A. R. MacDonald, G. Kreutz, & L. Mitchell (Eds.), *Music, health and wellbeing* (pp. 324–338). Oxford: Oxford University Press.

Schellenberg, E. G., Iverson, P., & McKinnon, M. C. (1999). Name that tune: Identifying popular recordings from brief excerpts. *Psychonomic Bulletin & Review*, *64*, 641–646.

Schellenberg, E. G., & Trehub, S. E. (1996). Natural musical intervals: Evidence from infant listeners. *Psychological Science*, *7*, 272–277.

Schellenberg, E. G., & Trehub, S. E. (2003). Good pitch memory is widespread. *Psychological Science*, *14*, 262–266.

Schellenberg, E. G., & von Scheve, C. (2012). Emotional cues in American popular music: Five decades of the Top 40. *Psychology of Aesthetics, Creativity, and the Arts*, *6*, 196–203.

Schenker, H. (1935). *Free composition* [*Der freie Satz*] (E. Oster, Trans.). New York: Longman.

Schenker, H. (1954). *Harmony* (E. M. Borgese, Trans.). Chicago, IL: Chicago University Press. (Original work published 1906)

Scherer, K. R., Banse, R., & Wallbott, H. G. (2001). Emotion inferences from vocal expression correlate across language and cultures. *Journal of Cross-Cultural Psychology*, *32*, 76–92.

Schlaug, G. (2009). Music, musicians, and brain plasticity. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford handbook of music psychology* (pp. 197–207). New York: Oxford University Press.

Schlaug, G., Jäncke, L., Huang, Y., & Steinmetz, H. (1995, February 3). In vivo evidence of structural brain asymmetry in musicians. *Science*, *267*, 699–701.

Schlaug, G., Jäncke, L., Huang, Y., Staiger, J. F., & Steinmetz, H. (1995). Increased corpus callosum size in musicians. *Neuropsychologia*, *33*, 1047–1055.

Schleuter, S. L. (1997). *A sound approach to teaching instrumentalists* (2nd ed.). Kent, OH: Kent State University Press.

Schmidt, C. P. (2005). Relations among motivation, performance achievement, and music experience variables in secondary instrumental music students. *Journal of Research in Music Education*, *53*, 134–147.

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review*, *82*, 225–260.

Schmuckler, M. A. (1999). Testing models of melodic contour similarity. *Music Perception*, *16*, 295–326.

Schmuckler, M. A. (2010). Melodic contour similarity using folk melodies. *Music Perception*, *28*, 169–194.

Schneider, P., Sluming, V., Roberts, N., Bleeck, S., & Rupp, A. (2005). Structural, functional, and perceptual differences in Heschyl's gyrus and musical instrument preference. *Annals of the New York Academy of Sciences*, *1060*, 387–394.

Schnupp, J., Nelken, I., & King, A. J. (2011). *Auditory neuroscience*. Cambridge, MA: MIT Press.

Schouten, J. F. (1940). The residue and the mechanism of hearing. *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen*, *43*, 991–999.

Schubert, E. (1996). Enjoyment of negative emotions in music. *Psychology of Music*, *24*, 18–28.

Schubert, E., & Stevens, C. (2006). The effect of implied harmony, contour and musical expertise on judgments of similarity of familiar melodies. *Journal of New Music Research*, *35*, 161–174.

Schulkin, J., & Raglan, G. B. (2014). The evolution of music and human social capability. *Frontiers in Human Neuroscience*, *8*, 292.

Schulkind, M. D., Posner, R. J., & Rubin, D. C. (2003). Musical features that facilitate melody identification: How do you know it's 'your' song when they finally play it? *Music Perception*, *21*, 217–249.

Schutz, M., & Kubovy, M. (2009). Deconstructing a musical illusion: Point-light representations capture salient properties of impact motions. *Canadian Acoustics*, *37*, 23–28.

Schutz, M., & Lipscomb, S. (2007). Hearing gestures, seeing music: Vision influences perceived tone duration. *Perception*, *36*, 888–897.

Schutz, M., & Vaisberg, J. (2014). Surveying the temporal structure of sounds used in music perception. *Music Perception*, *31*, 288–296.

Scruton, R. (1983). *The aesthetic understanding*. Manchester: Carcanet.

Seashore, C. E. (1919). *The psychology of musical talent*. Boston, MA: Silver Burdett.

Seashore, C. E. (1967). *The psychology of music*. New York: McGraw-Hill. (Original work published 1938)

Seddon, F. A., & O'Neill, S. A. (2001). An evaluation study of computer-based compositions by children with and without prior experience of formal instrumental music tuition. *Psychology of Music*, *29*, 4–19.

Seddon, F. A., & O'Neill, S. A. (2006). How does formal instrumental music tuition impact self- and teacher-evaluations of computer-based compositions? *Psychology of Music*, *34*, 27–45.

Seppänen, M., Brattico, E., & Tervaniemi, M. (2007). Practice strategies of musicians modulate neural processing and the learning of sound patterns. *Neurobiology of Learning and Memory*, *87*, 236–247.

Sergeant, D. C., & Himonides, E. (2016). Gender and music composition: A study of music, and the genderings of meanings. *Frontiers in Psychology*, *7*, 411.

Seuss, T. G. (1957). *The cat in the hat*. New York: Beginner Books/Random House.

Shenfield, T., Trehub, S. E., & Nakata, T. (2003). Maternal singing modulates infant arousal. *Psychology of Music*, *31*, 365–375.

Shepard, R. N. (1965). Approximation to uniform gradients of generalization by monotone transformations of scale. In D. I. Mostofsky (Ed.), *Stimulus generalization* (pp. 94–110). Stanford, CA: Stanford University Press.

Shepard, R. N. (1982). Structural representations of musical pitch. In D. Deutsch (Ed.), *The psychology of music* (pp. 344–390). New York: Academic Press.

Shevy, M., & Hung, K. (2013). Music in television advertising and other persuasive media. In S.-L. Tan, A. J. Cohen, S. D. Lipscomb, & R. A. Kendall (Eds.), *The psychology of music in multimedia* (pp. 315–338). Oxford: Oxford University Press.

Shoda, H., & Adachi, M. (2015). Why live recording sounds better: A case study of Schumann's *Träumerei*. *Frontiers in Psychology*, *5*, 1564.

Shuter-Dyson, R., & Gabriel, C. (1981). *The psychology of musical ability*. London: Methuen.

Simoens, V. L., Puttonen, S., & Tervaniemi, M. (2015). Are music performance anxiety and performance boost perceived as extremes of the same continuum? *Psychology of Music*, *43*, 171–187.

Simoens, V. L., & Tervaniemi, M. (2013). Musician–instrument relationship as a candidate index for professional well-being in musicians. *Psychology of Aesthetics, Creativity, and the Arts*, *7*, 171–180.

Simon, H. A., & Chase, W. G. (1973). Skill in chess. *American Scientist*, *61*, 394–403.

Simonton, D. K. (1991). Emergence and realization of genius: The lives and works of 120 classical composers. *Journal of Personality and Social Psychology*, *61*, 829–840.

Simonton, D. K. (2014). Creative performance, expertise acquisition, individual differences, and developmental antecedents: An integrative research agenda. *Intelligence*, *45*, 66–73.

Sloboda, J. A. (1974). The eye–hand span – an approach to the study of sight reading. *Psychology of Music*, *2*, 4–10.

Sloboda, J. A. (1976). The effect of item position on the likelihood of identification by inference in prose reading and music reading. *Canadian Journal of Experimental Psychology*, *30*, 228–237.

Sloboda, J. A. (1985). *The musical mind: The cognitive psychology of music*. Oxford: Oxford University Press.

Sloboda, J. A. (1991). Music structure and emotional response: Some empirical findings. *Psychology of Music*, *19*, 110–120.

Sloboda, J. A. (2005). *Exploring the musical mind: Cognition, emotion, ability, function*. Oxford: Oxford University Press.

Sloboda, J. A., Davidson, J. W., Howe, M. J. A., & Moore, D. G. (1996). The role of practice in the development of performing musicians. *British Journal of Psychology*, *87*, 287–309.

Sloboda, J. A., & O'Neill, S. A. (2001). Emotions in everyday listening to music. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 415–430). Oxford: Oxford University Press.

Sloboda, J. A., Wise, K. J., & Peretz, I. (2005). Quantifying tone deafness in the general population. *Annals of the New York Academy of Sciences*, *1060*, 255–261.

Smith, J. D., Kemler Nelson, D. G., Grohskopf, L. A., & Appleton, T. (1994). What child is this? What interval was that? Familiar tunes and music perception in novice listeners. *Cognition*, *52*, 23–54.

Snyder, B. (2000). *Music and memory: An introduction*. Cambridge, MA: MIT Press.
Sole, M. (2017). Crib song: Insights into functions of toddlers' private spontaneous singing. *Psychology of Music*, *45,* 172–192.
Soley, G., & Hannon, E. E. (2010). Infants prefer the musical meter of their own culture: A cross-cultural comparison. *Developmental Psychology*, *46*, 286–292.
Sontag, L. W., & Wallace, R. F. (1935). The movement response of the human fetus to sound stimuli. *Child Development*, *6*, 253–258.
Sosniak, L. A. (1985). Learning to be a concert pianist. In B. S. Bloom (Ed.), *Developing talent in young people* (pp. 19–67). New York: Ballantine.
Sowiński, J., & Dalla Bella, S. (2013). Poor synchronization to the beat may result from deficient auditory–motor mapping. *Neuropsychologia*, *51*, 1952–1963.
Spahn, C., Wasmer, C., Eickhoff, F., & Nusseck, M. (2014). Comparing violinists' body movements while standing, sitting, and in sitting orientations to the right or left of a music stand. *Medical Problems of Performing Artists*, *29*, 86–93.
Spielberg, S. (Director), Goldman, G., Shusett, R. (Executive Producers), & Williams, J. (Original score). (2002). *Minority report* [Motion picture]. United States: DreamWorks LLC and Twentieth Century Fox.
St. John, P. A. (2006). Finding and making meaning: Young children as musical collaborators. *Psychology of Music*, *34*, 238–261.
Standley, J. M., & Walworth, D. (2010). *Music therapy with premature infants: Research and developmental interventions* (2nd ed.). Silver Spring, MD: American Music Therapy Association.
Stevens, C., & Byron, T. (2016). Universals in music processing: Entrainment, acquiring expectations, and learning. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford handbook of music psychology* (2nd ed., pp. 19–32). Oxford: Oxford University Press.
Stickgold, R., & Walker, M. P. (2005). Memory consolidation and reconsolidation: What is the role of sleep? *Trends in Neuroscience*, *28*, 408–415.
Stone, R. M. (1985). In search of time in African music. *Music Theory Spectrum*, *7*, 139–148.
Sugimoto, T., Kobayashi, H., Nobuyoshi, N., Kiriyama, Y., Takeshita, H., Nakamura, T., & Hashiya, K. (2010). Preference for consonant music over dissonant music by an infant chimpanzee. *Primates*, *51*, 7–12.
Sundberg, J. (1982). Perception of singing. In D. Deutsch (Ed.), *The psychology of music* (pp. 59–98). New York: Academic Press.
Sundberg, J. (1987). *The science of the singing voice*. Dekalb, IL: Northern Illinois University Press.
Suomen Akatemia (2011, December 6). Listening to music lights up the whole brain. *ScienceDaily*. Retrieved from https://www.sciencedaily.com/releases/2011/12/111205081731.htm
Swaminathan, S., & Schellenberg, E. G. (2015). Current emotion research in music psychology. *Emotion Review*, *7*, 189–197.
Tafuri, J. (2008). *Infant musicality.* Farnham: Ashgate.
Tahlier, M., Miron, A. M., & Rauscher, F. H. (2013). Music choice as a sadness regulation strategy for resolved versus unresolved sad events. *Psychology of Music*, *41*, 729–748.
Takeuchi, A. H., & Hulse, S. H. (1993). Absolute pitch. *Psychological Bulletin*, *113*, 345–361.
Tan, S.-L. (2002). Beginners' intuitions about musical notation. *College Music Symposium*, *42*, 131–141.
Tan, S.-L. (2017a). From intuition to evidence: The experimental psychology of film music. In M. Mera, R. A. Sadoff, & B. Winters (Eds.), *The Routledge companion to screen music and sound* (pp. 517–530). New York: Routledge.
Tan, S.-L. (2017b). Scene and heard: The role of music in shaping interpretations of film. In R. Ashley & R. Timmers (Eds.), *The Routledge companion to music cognition* (pp. 363–376). New York: Routledge.
Tan, S.-L., Cohen, A. J., Lipscomb, S. D., & Kendall, R. A. (Eds.). (2013). *The psychology of music in multimedia*. Oxford: Oxford University Press.

Tan, S.-L., Spackman, M. P., & Bezdek, M. A. (2007). Viewers' interpretations of film characters' emotions: Effects of presenting film music before or after a character is shown. *Music Perception*, *25*, 135–152.

Tan, S.-L., Spackman, M. P., & Wakefield, E. M. (2017). The effects of diegetic and nondiegetic music on viewers' interpretations of a film scene. *Music Perception*, *34*, 605–623.

Tan, S.-L., Wakefield, E. M., & Jeffries, P. W. (2009). Musically untrained college students' interpretations of musical notation: Sound, silence, loudness, duration, and temporal order. *Psychology of Music*, *37*, 5–24.

Tan, Y. T., McPherson, G. E., Peretz, I., Berkovic, S. F., & Wilson, S. J. (2014). The genetic basis of music ability. *Frontiers in Psychology*, *5*, 658.

Tarnowski, S. M. (1999). Musical play and young children. *Music Educators Journal*, *86*, 26–29.

Taruffi, L., & Koelsch, S. (2014). The paradox of music-evoked sadness: An online survey. *PLOS ONE*, *9*, e110490.

Teki, S., Grube, M., Kumar, S., & Griffiths, T. D. (2011). Distinct neural substrates of duration-based and beat-based auditory timing. *Journal of Neuroscience*, *31*, 3805–3812.

Temperley, D. (2000). Meter and grouping in African music: A view from music theory. *Ethnomusicology*, *44*, 65–96.

Temperley, D. (2001). *The cognition of basic musical structures*. Cambridge, MA: MIT Press.

Tenzer, M. (2000). *Gamelan gong kebyar*. Chicago, IL: University of Chicago Press.

Terhardt, E. (1974). Pitch, consonance, and harmony. *Journal of the Acoustical Society of America*, *55*, 1061–1069.

Terwogt, M. M., & van Grinsven, F. (1991). Musical expression of moodstates. *Psychology of Music*, *19*, 99–109.

Thayer, J. F., & Levenson, R. W. (1983). Effects of music on psychophysiological responses to a stressful film. *Psychomusicology*, *3*, 44–52.

Theiler, A. M., & Lippman, L. G. (1995). Effects of mental practice and modeling on guitar and vocal performance. *Journal of General Psychology*, *122*, 329–343.

Theusch, E., Basu, A., & Gitschier, J. (2009). Genome-wide study of families with absolute pitch reveals linkage to 8q24.21 and locus heterogeneity. *American Journal of Human Genetics*, *85*, 112–119.

Thomas, J. P., & Nettelbeck, T. (2014). Performance anxiety in adolescent musicians. *Psychology of Music*, *42*, 624–634.

Thompson, W. F., & Balkwill, L. L. (2010). Cross-cultural similarities and differences. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 755–788). Oxford: Oxford University Press.

Thompson, W. F., Cuddy, L., & Plaus, C. (1997). Expectancies generated by melodic intervals: Evaluation of principles of melodic implication in a melody-completion task. *Perception & Psychophysics*, *59*, 1069–1076.

Thompson, W. F., Graham, P., & Russo, F. A. (2005). Seeing music performance: Visual influences on perception and experience. *Semiotica*, *156*, 203–227.

Thompson, W. F., Russo, F. A., & Livingstone, S. R. (2010). Facial expressions of singers influence perceived pitch relations. *Psychonomic Bulletin Review*, *17*, 317–322.

Thompson, W. F., Russo, F. A., & Sinclair, D. (1994). Effects of underscoring on the perception of closure in filmed events. *Psychomusicology*, *13*, 9–27.

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2001). Arousal, mood, and the Mozart effect. *Psychological Science*, *12*, 248–251.

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2004). Decoding speech prosody: Do music lessons help? *Emotion*, *4*, 46–64.

Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review*, *107*, 885–913.

Tillmann, B., & Bigand, E. (2002). A comparative review of priming effects in language and music. In P. McKevitt, S. Ó. Nulláin, & C. Mulvihill (Eds.), *Language, vision, and music: Selected papers from the 8th International Workshop on the Cognitive Science of Natural Language Processing,* *Galway, Ireland 1999. Advances in consciousness research* (pp. 231–240). Amsterdam: John Benjamins.

Tillmann, B., Janata, P., & Bharucha, J. J. (2003). Activation of the inferior frontal cortex in musical priming. *Cognitive Brain Research*, *16*, 145–161.

Timmers, R., & Ashley, R. (2007). Emotional ornamentation in performances of a Handel sonata. *Music Perception*, *25*, 117–134.

Timmers, R., Endo, S., Bradbury, A., & Wing, A. M. (2014). Synchronization and leadership in string quartet performance: a case study of auditory and visual cues. *Frontiers in Psychology*, *5*, article 645.

Todd, N. (1985). A model of expressive timing in music performance in tonal music. *Music Perception*, *3*, 33–58.

Todd, N. P. M., & Lee, C. S. (2015). The sensory-motor theory of rhythm and beat induction 20 years on: A new synthesis and future perspectives. *Frontiers in Human Neuroscience*, *9*, 444.

Trainor, L. J. (1996). Infant preferences for infant-directed versus noninfant-directed playsongs and lullabies. *Infant Behavior & Development*, *19*, 83–92.

Trainor, L. J., Lee, K., & Bosnyak, D. J. (2011). Cortical plasticity in 4-month-old infants: Specific effects of experience with musical timbres. *Brain Topography*, *24*, 192–203.

Trainor, L. J., & Trehub, S. E. (1994). Key membership and implied harmony in Western tonal music: Developmental perspectives. *Perception & Psychophysics*, *56*, 125–132.

Trainor, L. J., Tsang, C. D., & Cheung, V. H. W. (2002). Preference for sensory consonance in 2- and 4-month-old infants. *Music Perception*, *20*, 187–194.

Trainor, L. J., Wu, L., & Tsang, C. D. (2004). Long-term memory for music: Infants remember tempo and timbre. *Developmental Science*, *7*, 289–296.

Tramo, M. J., Lense, M., van Ness, C., Kagan, J., Settle, M. D., & Cronin, J. H. (2011). Effects of music on physiological and behavioral indices of acute pain and stress in premature infants: Clinical trial and literature review. *Music and Medicine*, *3*, 72–83.

Trehub, S. E. (1993). Temporal auditory processing in infancy. *Annals of the New York Academy of Sciences*, *682*, 137–149.

Trehub, S. (2000). Human processing predispositions and musical universals. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 427–448). Cambridge, MA: MIT Press.

Trehub, S. E., Bull, D., & Thorpe, L. A. (1984). Infants' perception of melodies: The role of melodic contour. *Child Development*, *55*, 821–830.

Trehub, S. E., & Gudmundsdottir, H. R. (2015). Mothers as singing mentors for infants. In G. Welch, D. Howard, & J. Nix (Eds.), *The Oxford handbook of singing*. Oxford: Oxford University Press.

Trehub, S. E., Hill, D. S., & Kamenetsky, S. B. (1997). Parents' sung performances for infants. *Canadian Journal of Experimental Psychology*, *51*, 385–396.

Trehub, S. E., Schellenberg, E. G., & Kamenetsky, S. B. (1999). Infants' and adults' perception of scale structure. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 965–975.

Trehub, S. E., & Thorpe, L. A. (1989). Infants' perception of rhythm: Categorization of auditory sequences by temporal structure. *Canadian Journal of Psychology*, *43*, 217–229.

Trehub, S. E., Thorpe, L. A., & Morrongiello, B. A. (1985). Infants' perception of melodies: Changes in a single tone. *Infant Behavior & Development*, *8*, 213–223.

Trehub, S. E., Unyk, A. M., & Henderson, J. L. (1994). Children's songs to infant siblings: Parallels with speech. *Journal of Child Language*, *21*, 735–744.

Triplett, N. (1898). The dynamogenic factors in pacemaking and competition. *American Journal of Psychology*, *9*, 507–533.

Trout, J. D. (2001). The biological basis of speech: What to infer from talking to the animals. *Psychological Review*, *108*, 523–549.

Urmson, J. O. (1973). Representation in music. In G. Vesey (Ed.), *Philosophy and the arts, Royal Institute of Philosophy lectures* (Vol. 6, pp. 132–146). London: Macmillan.

Valentine, C. W. (1962). Musical intervals and attitudes to music. In C. W. Valentine (Ed.), *The experimental psychology of beauty* (pp. 196–227). London: Methuen.

van den Bergh, B. R. H. (1992). Maternal emotions during pregnancy and fetal and neonatal behavior. In J. G. Nijhuis (Ed.), *Fetal behavior: Developmental and perinatal aspects* (pp. 157–208). Oxford: Oxford University Press.

Van Hedger, S. C., Hogstrom, A., Palmer, C., & Nusbaum, H. C. (2015). Sleep consolidation of musical competence. *Music Perception*, 33, 163–178.

van Puyvelde, M., Loots, G., Vinck, B., De Coster, L., Matthijs, L., Mouvet, K., & Pattyn, N. (2013). The interplay between tonal synchrony and social engagement in mother–infant interaction. *Infancy*, *18*, 849–872.

van Puyvelde, M., Vanfleteren, P., Loots, G., Deschuyffeleer, S., Vinck, B., Jacquet, W., & Verhelst, W. (2010). Tonal synchrony in mother-infant interaction based on harmonic and pentatonic series. *Infant Behavior & Development*, *33*, 387–400.

van Vugt, F. T., Jabusch, H.-C., & Altenmüller, E. (2012). Fingers phrase music differently: Trial-to-trial variability in piano scale playing and auditory perception reveal motor chunking. *Frontiers in Psychology*, *3*, 495.

van Zijl, A. G. W., Toiviainen, P., Lartillot, O., & Luck, G. (2014). The sound of emotion: The effect of performers' experienced emotions on auditory performance characteristics. *Music Perception*, 32, 33–50.

Varvarigou, M., & Green, L. (2014). Music 'learning styles' and 'learning strategies' in the instrumental lesson: The Ear Playing Project (EPP). *Psychology of Music*, *43*, 705–422.

Veneklasen, P. S. (1975). Design considerations from the viewpoint of the professional consultant. In R. Mackenzie (Ed.), *Auditorium acoustics: The proceedings of an international symposium on architectural acoustics*. New York: Wiley-Halsted.

Vieillard, S., & Gilet, A.-L. (2013). Age-related differences in affective responses to and memory for emotions conveyed by music: A cross-sectional study. *Frontiers in Psychology*, *4*, 711.

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., & Levitin, D. J. (2006). Cross-modal interactions in the perception of musical performance. *Cognition*, *101*, 80–113.

Vitouch, O. (2001). When your ear sets the stage: Musical context effects in film perception. *Psychology of Music*, *29*, 70–83.

Volkova, A., Trehub, S. E., & Schellenberg, E. G. (2006). Infants' memory for musical performances. *Developmental Science*, *9,* 583–589.

Vongpaisal, T., Trehub, S. E., & Schellenberg, E. G. (2006). Song recognition by children and adolescents with cochlear implants. *Journal of Speech, Language, and Hearing Research*, *49*, 1091–1103.

von Hippel, P. & Huron, D. (2000). Why do skips precede reversals? The effects of tessitura and melodic structure. *Music Perception*, *18*, 59–86.

Vuoskoski, J. K., Thompson, W. F., McIlwain, D., & Eerola, T. (2012). Who enjoys listening to sad music and why? *Music Perception*, *29*, 311–317.

Vygotsky, L. S. (1978). *Mind in society: The development of higher mental processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Cambridge, MA: Harvard University Press. (Original work published 1930)

Walker, A. (1983). *Franz Liszt: The virtuoso years 1811–1847*. New York: Cornell University Press.

Warren, J. D., Uppenkamp, S., Patterson, R. D., & Griffiths, T. D. (2003). Separating pitch chroma and pitch height in the human brain. *Proceedings of the National Academy of Sciences*, *100*, 10038–10042.

Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science*, *167*, 392–393.

Warren, R. M. (1999). *Auditory perception: A new analysis and synthesis*. Cambridge: Cambridge University Press.

Webb, A. R., Heller, H. T., Benson, C. B., & Lahav, A. (2015). Mother's voice and heartbeat sounds elicit auditory plasticity in the human brain before full gestation. *Proceedings of the National Academy of Sciences*, *112*, 3152–3157.

Wehr-Flowers, E. (2006). Differences between male and female students' confidence, anxiety, and attitude toward learning jazz improvisation. *Journal of Research in Music Education*, *54*, 337–349.

Welch, G. F. (1979). Poor pitch singing: A review of the literature. *Psychology of Music*, *7*, 50–58.

Welch, G. F. (1998). Early childhood musical development. *Research Studies in Music Education*, *11*, 27–41.

Welch, G. F., Himonides, E., Saunders, J., Papageorgi, I., & Sarazin, M. (2014). Singing and social inclusion. *Frontiers in Psychology*, *5*, 803.

Werker, J. F., Pegg, J. E., & McLeod, P. J. (1994). A cross-language investigation of infant preference for infant-directed communication. *Infant Behavior & Development*, *17*, 323–333.

Wiggins, J., & Espeland, M. (2012). Creating in music learning contexts. In G. McPherson & G. Welch (Eds.), *Oxford handbook of music education* (pp. 341–360). New York: Oxford University Press.

Wilhelm, K., Gillis, I., Schubert, E., & Whittle, E. L. (2013). On a blue note: Depressed peoples' reasons for listening to music. *Music and Medicine*, *5*, 76–83.

Wilkin, P. (1996). A comparison of fetal and newborn responses to music and sound stimuli with and without daily exposure to a specific piece of music. *Bulletin of the Council for Research in Music Education*, *127*, 163–169.

Williamon, A., Aufegger, L., & Eiholzer, H. (2014). Simulating and stimulating performance: Introducing distributed simulation to enhance musical learning and performance. *Frontiers of Psychology*, *5*, 25.

Williamon, A., & Valentine, E. (2000). Quantity and quality of musical practice as predictors of performance quality. *British Journal of Psychology*, *91*, 353–376.

Williamon, A., & Valentine, E. (2002). The role of retrieval structures in memorizing music. *Cognitive Psychology*, *44*, 1–32.

Williamson, V., Liikkanen, L., Jakubowski, K., & Stewart, L. (2014). Sticky tunes: How do people react to involuntary musical imagery? *PLOS ONE*, *9*, e86170.

Williamson, V. J., Jilka, S. R., Fry, J., Finkel, S., Müllensiefen, D., & Stewart, L. (2011). How do 'earworms' start? Classifying the everyday circumstances of involuntary musical imagery. *Psychology of Music*, *40*, 259–284.

Wiltermuth, S. S., & Heath, C. (2009). Synchrony and cooperation. *Psychological Science*, *20*, 1–5.

Wing, A. M., & Kristofferson, A. B. (1973). Response delays and the timing of discrete motor responses. *Perception & Psychophysics*, *14*, 5–12.

Winkler, I., Háden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 2468–2471.

Winner, E. (2000). The origins and ends of giftedness. *American Psychologist*, *55*, 159–169.

Witek, M. A. G., Clarke, E. F., Wallentin, M., Kringelbach, M. L., & Vuust, P. (2014). Syncopation, body-movement and pleasure in groove music. *PLOS ONE*, *9*, e94446.

Wöllner, C., & Auhagen, W. (2008). Perceiving conductors' expressive gestures from different visual perspectives. *Music Perception*, *26*, 129–143.

Wöllner, C., & Halpern, A. R. (2016). Attentional flexibility and memory capacity in conductors and pianists. *Attention, Perception & Psychophysics*, *78*, 198–208.

Wong, P. C. M., Parsons, L. M., Martinez, M., & Diehl, R. L. (2004). The role of the insular cortex in pitch pattern perception: The effect of linguistic contexts. *Journal of Neuroscience*, *24*, 9153–9160.

Wong, P. C. M., Roy, A. K., & Margulis, E. H. (2009). Bimusicalism: The implicit dual enculturation of cognitive and affective systems. *Music Perception*, *27*, 81–88.

Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, *10*, 420–422.

Wood, D., Bruner, J., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, *17*, 89–100.

Woodward, S. C. (1992). Intrauterine rhythm and blues? *British Journal of Obstetrics and Gynaecology*, *99*, 787–790.

Woody, R. H., & Lehmann, A. C. (2010). Student musicians' ear-playing ability as a function of vernacular music experiences. *Journal of Research in Music Education*, *58*, 101–115.

Wright, A. A., Rivera, J. J., Hulse, S. H., Shyan, M., & Neiworth, J. J. (2000). Music perception and octave generalization in rhesus monkeys. *Journal of Experimental Psychology: General*, *129*, 291–307.

Yeoh, J. P. S., & North, A. C. (2010). The effect of musical fit on consumers' memory. *Psychology of Music*, *38*, 368–378.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, *18*, 459–482.

Yost, W. A. (2000). *Fundamentals of hearing* (4th ed.). San Diego, CA: Academic Press.

Young, S. (2003a). The interpersonal dimension: A potential source of musical creativity for young children? *Musicae Scientiae, Special Issue: Musical Creativity Special 10th Anniversary Conference Issue: Award Papers*, *2003–2004*, 175–191.

Young, S. (2003b). Time-space structuring in spontaneous play on educational percussion instruments among three- and four-year-olds. *British Journal of Music Education*, *20*, 45–59.

Zajonc, R. B. (1965). Social facilitation. *Science*, *149*, 261–274.

Zamm, A., Pfordresher, P. Q., & Palmer, C. (2015). Temporal coordination in joint music performance: Effects of endogenous rhythms and auditory feedback. *Experimental Brain Research*, *233*, 607–615.

Zamm, A., Wellman, C., & Palmer, C. (2016). Endogenous rhythms influence interpersonal synchrony. *Journal of Experimental Psychology: Human Perception & Performance*, *42*, 161–166.

Zarate, J. M. (2013). The neural control of singing. *Frontiers in Human Neuroscience*, *7*, 237

Zarco, W., Merchant, H., Prado, L., & Mendez, J. C. (2009). Subsecond timing in primates: Comparison of interval production between human subjects and rhesus monkeys. *Journal of Neurophysiology*, *102*, 3191–3202.

Zatorre, R. J., & Baum, S. R. (2012). Musical melody and speech intonation: Singing a different tune. *PLOS Biology*, *10*, e1001372.

Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences*, *6*, 37–46.

Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: Auditory–motor interactions in music perception and production. *Nature Reviews Neuroscience*, *8*, 547–558.

Zatorre, R. J., Evans, A. C., Meyer, E., & Gjedde, A. (1992, May 8). Lateralization of phonetic and pitch discrimination in speech processing. *Science*, *256*, 846–849.

Zatorre, R., J. & Halpern, A. (2005). Mental concerts: Musical imagery and the auditory cortex. *Neuron*, *47*, 9–12.

Zatorre, R. J., Halpern, A. R., Perry, D. W., Meyer, E., & Evans, A. C. (1995). Hearing in the mind's ear: A PET investigation of musical imagery and perception. *Journal of Cognitive Neuroscience*, *8*, 29–46.

Zaza, C. (1993). Prevention of musicians' playing-related health problems: Rationale and recommendations for action. *Medical Problems of Performing Artists*, *8*, 117–121.

Zentner, M., & Eerola, T. (2010). Rhythmic engagement with music in infancy. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 5768–5773.

Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion*, *8*, 494–521.

Zimmer, E. Z., Divon, M. Y., Vilensky, A., Sarna, Z., Peretz, B. A., & Paldi, E. (1982). Maternal exposure to music and fetal activity. *European Journal of Obstetrics and Gynecology and Reproductive Biology*, *13*, 209–213.

Zimmerman, E., Keunen, K., Norton, M., & Lahav, A. (2013). Weight gain velocity in very low-birth-weight infants: Effects of exposure to biological maternal sounds. *American Journal of Perinatology*, *30*, 863–870.

Name index

Subject index