sempre:

Society for Education, Music and Psychology Research

ROUTLEDGE

Music and Mental Imagery

Edited by Mats B. Küssner, Liila Taruffi and Georgia A. Floridou

SEMPRE Studies in The Psychology of Music

Music and Mental Imagery

Drawing on perspectives from music psychology, cognitive neuroscience, philosophy, musicology, clinical psychology, and music education, *Music and Mental Imagery* provides a critical overview of cutting-edge research on the various types of mental imagery associated with music. The four main parts cover an introduction to the different types of mental imagery associated with music such as auditory/musical, visual, kinaesthetic, and multimodal mental imagery; a critical assessment of established and novel ways to measure mental imagery in various musical contexts; coverage of different states of consciousness, all of which are relevant for, and often associated with, mental imagery in music, and a critical overview of applications of mental imagery in health, educational, and performance settings.

By both critically reviewing up-to-date scientific research and offering new empirical results, this book provides a unique overview of the different types and origins of mental imagery in musical contexts, various ways to measure them, and intriguing insights into related mental phenomena such as mind-wandering and synaesthesia. This will be of particular interest for scholars and researchers of music psychology and music education. It will also be useful for practitioners working with music in applied health and educational contexts.

Mats B. Küssner is Lecturer in the Department of Musicology and Media Studies at Humboldt-Universität zu Berlin and Visiting Research Fellow in the Department of Psychology at Goldsmiths, University of London.

Liila Taruffi is Lecturer in Music Psychology at Durham University, UK. She has an interdisciplinary background in psychology, neuroscience, and aesthetics.

Georgia A. Floridou is a psychologist, researcher, and educator operating at the intersection of music, psychology, and neuroscience. She is Honorary Research Fellow in the Department of Music at the University of Sheffield, UK.

SEMPRE Studies in The Psychology of Music

Series Editors
Graham F. Welch
UCL Institute of Education, University College London, UK
Adam Ockelford
University of Roehampton, UK
Ian Cross
University of Cambridge, UK

The theme for the series is the psychology of music, broadly defined. Topics include (i) musical development at different ages, (ii) exceptional musical development in the context of special educational needs, (iii) musical cognition and context, (iv) culture, mind and music, (v) micro to macro perspectives on the impact of music on the individual (from neurological studies through to social psychology), (vi) the development of advanced performance skills, and (vii) affective perspectives on musical learning. The series presents the implications of research findings for a wide readership, including user-groups (music teachers, policy makers, and parents) as well as the international academic and research communities. This expansive embrace, in terms of both subject matter and intended audience (drawing on basic and applied research from across the globe), is the distinguishing feature of the series, and it serves SEMPRE's distinctive mission, which is to promote and ensure coherent and symbiotic links between education, music and psychology research.

The Artist and Academia
Edited by Helen Phelan and Graham F. Welch

Expanding Professionalism in Music and Higher Music Education
A Changing Game
Edited by Helena Gaunt and Heidi Westerlund

Indian Classical Music and the Gramophone, 1900–1930
Vikram Sampath

Body and Force in Music
Metaphoric Constructions in Music Psychology
Youn Kim

Music and Mental Imagery
Edited by Mats B. Küssner, Liila Taruffi and Georgia A. Floridou

For more information about this series, please visit: www.routledge.com/music/series/SEMPRE

Music and Mental Imagery

Edited by Mats B. Küssner,
Liila Taruffi and Georgia A. Floridou

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

Cover image: © Getty images

First published 2023
by Routledge
4 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

and by Routledge
605 Third Avenue, New York, NY 10158

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

This work was generously supported by the Chair of Transcultural Musicology and Historical Anthropology of Music in the Department of Musicology and Media Studies at Humboldt-Universität zu Berlin.

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

ISBN: 978-0-367-35216-5 (hbk)
ISBN: 978-1-032-37607-3 (pbk)
ISBN: 978-0-429-33007-0 (ebk)

DOI: 10.4324/9780429330070

Typeset in Times New Roman
by Apex CoVantage, LLC

Access the Support Material: www.routledge.com/9780367352165

Contents

List of Figures, Tables and eResources

Figures

Tables

eResources

The eResources can be accessed at www.routledge.com/9780367352165.

Tables

Chapter

Contributors

George Athanasopoulos received the PhD degree from University of Edinburgh in 2013. He is a researcher in cognitive ethnomusicology, focusing on the cross-cultural perception of music. He has conducted fieldwork research in Great Britain, Greece, Japan, northwest Pakistan, and Papua New Guinea. His work has been funded by the Marie Curie Foundation/COFUND, the Great Britain Sasakawa Foundation, the University of Edinburgh, the Aristotle University of Thessaloniki (State Scholarships Foundation), and the Alexander S. Onassis Foundation. His research has been published in *PLOS ONE*, the *Annals of the New York Academy of Sciences*, *Psychology of Music*, *Musicae Scientiae*, and *Empirical Musicology Review*.

Amy M. Belfi is Assistant Professor in the Department of Psychological Science at Missouri University of Science and Technology. She received her BA degree in psychology from St. Olaf College, her PhD degree in neuroscience from the University of Iowa, and completed postdoctoral training at New York University. Her work covers a broad range of topics in the field of music cognition, such as music and autobiographical memory, musical imagery, aesthetic judgements of music, and musical anhedonia. She uses a variety of techniques to investigate these subjects, including behavioural studies, functional neuroimaging, psychophysiology, and neuropsychological approaches.

Dr Mary T. Black is a singer, lecturer, presenter, and conductor whose long-held interest in choral singing and directing led her to complete a PhD degree on the functions and effects of imagery in choral directing at Leeds University. She has presented at conferences on this topic both nationally and internationally, most recently in Lund (Sweden), Truro, Berlin, York, Birmingham, and Graz (Austria). Her most recent publication is Black, M. T. (May 2020). “That Was a Bit of a Stodgy Start!” The functions and effects of verbal imagery in choral rehearsals. (M. Ashley, Ed.) *Choral Directions Research Journal*, 1(1), 1–16. Mary is Visiting Research Fellow in the School of Music at Leeds University and is Music Director of Liverpool Phoenix Voices.

Robina A. Day completed her PhD research on the relationship between music-evoked emotions and visual imagery experiences at the Music, Sound and

Performance Lab at Macquarie University. Her academic publications focus on the relationship between emotional experience and musical imagery.

Helena Dukić holds a BA degree in piano from Music Academy Zagreb (2010) and a BA degree in music from University of Cambridge in 2011. She studied film composition as her master's degree at Kingston University, London, in 2013, researching the influence of film music on the perception of image and movement on screen. Dukić is a Level III Guided Imagery and Music trainee and is currently in education for a Gestalt psychotherapist. She completed her PhD degree in 2019 at the Centre for Systematic Musicology, University of Graz, focusing on the narrative nature of music and imagery in GIM and has since presented her research in numerous international conferences and in four peer-reviewed papers.

Tuomas Eerola is Professor of Music Cognition at Durham University, UK. His disciplinary orientation is music psychology. Eerola has published more than 180 papers and book chapters on topics including musical similarity, melodic expectations, perception of rhythm and timbre, music-induced movement, consonance, and induction and perception of emotions. He is also on the editorial board of *Psychology of Music*, *Psychomusicology*, and *Frontiers in Psychology—Auditory Cognitive Neuroscience* and is a consulting editor for *Musicae Scientiae* and *Music Perception*.

Dr Jörg Fachner is Professor for Music, Health and the Brain and Co-Director of the Cambridge Institute for Music Therapy Research at Anglia Ruskin University, Cambridge, UK. He is researching music and consciousness states, music therapy and treatment of addiction, depression, and neurorehabilitation; currently, he is investigating hyper-scanning of music therapy.

Katherine K. Finch is currently completing her PhD degree in clinical psychology at the University of Waterloo and is a member of the Psychological Intervention Research Team, studying under the supervision of Dr Jonathan M. Oakman. Her clinical and research interests focus on imagery interventions to help musicians manage music performance anxiety.

Dr Georgia A. Floridou is a psychologist, researcher, and educator operating at the intersection of music, psychology, and neuroscience. She is Honorary Research Fellow in the Department of Music at the University of Sheffield, UK. Her interests focus on music perception and cognition, memory, mental imagery, spontaneous thoughts (e.g., mind-wandering), and how to harness them in health, well-being, and technology. The methods she uses include surveys, behavioural experiments, and neuroimaging techniques (EEG) and interviews and experience sampling. Georgia has studied and worked for psychology departments and music departments in Greece, the UK, and the Netherlands and has received funding from a variety of organizations, including the British Academy (UK), SEMPRE (UK), and IKY (Greece).

Rebecca W. Gelding completed her PhD research in the neural correlates of musical pitch and rhythm imagery. Her academic publications focus on pitch imagery, brain oscillations during music perception, and mathematical modelling of brain oscillations. She has written popular science articles for *Times Higher Education*, and appeared on ABC Classic and the All in the Mind podcast. Her TEDx Macquarie University 2019 presentation "Unlocking the Music in your Mind" was selected to be featured on ted.com.

Rolf Inge Godøy is a professor of music theory in the Department of Musicology, University of Oslo, and a researcher at this university's RITMO Centre for Interdisciplinary Studies in Rhythm, Time and Motion. His main interest is in phenomenological approaches to music theory, meaning taking our subjective experiences of music as the point of departure for music theory. This work has been expanded to include research on music-related body motion in performance and listening, using various conceptual and technological tools to explore the relationships between sound and body motion in the experience of music.

Anthony Gritten is Head of Undergraduate Programmes at the Royal Academy of Music in London. He has published articles on Bakhtin, Cage, Delius, Lyotard, Nancy, and Stravinsky, written for visual artists' catalogues and philosophy dictionaries, edited two volumes on music and gesture, and written about numerous issues in Performance Studies, including artistic research, collaboration, distraction, empathy, ensemble interaction, entropy, ergonomics, listening, problem-solving, recording, and timbre.

Andrea Halpern is Professor of Psychology at Bucknell University. She studies cognition and neuroscience of nonverbal information including auditory imagery, music perception, and emotion, including in older adults. She has also been researching individual differences in vocal pitch-matching as well some audience engagement projects. She has received grants from the US National Science Foundation, National Insitute on Aging, and Grammy Foundation, and served as President of the Society for Music Perception and Cognition. In her nonacademic life, she is a choral and chamber music singer, with a special interest in Early Music.

Dr Ruth Herbert is Director of Graduate Studies and Senior Lecturer in Music in the Department of Music and Audio Technology, University of Kent, UK. She is a music psychologist and performer who has published in the fields of music in everyday life, music, health and well-being, consciousness studies, sonic studies, evolutionary psychology, and music education. Ruth is the author of *Everyday Music Listening: Absorption, Dissociation and Trancing* (Routledge, 2011/2016) and the principal editor (with Eric Clarke & David Clarke) of *Music and Consciousness 2: Worlds, Practices, Modalities* (OUP, 2019). She is an editorial board member for the *Journal of Sonic Studies* and a consulting editor for *Musicae Scientiae*.

Anri Herbst is an associate professor at the South African College of Music, University of Cape Town, and the head of music education. As a DAAD scholarship holder, she researched aural perception for her PhD degree in 1993. Not only has she led several projects sponsored by the National Research Foundation in South Africa and Harry Oppenheimer Memorial Trust, but she also founded the ISI-accredited *Journal of the Musical Arts in Africa* as Editor-in-Chief.

Timothy L. Hubbard is an adjunct faculty member at Arizona State University and Grand Canyon University. He earned a PhD degree from Dartmouth College and a BA degree from University of Denver. He has published one book, more than 90 papers in peer-reviewed journals, 12 chapters in edited books, and made more than 100 scientific presentations. His research involves spatial representation, music cognition, and mental imagery. He is Associate Editor at *Auditory Perception & Cognition* and *Frontiers in Psychology* and Consulting Editor at *Attention, Perception, & Psychophysics* and *Journal of Experimental Psychology: Human Perception and Performance*. He is a fellow of the Association for Psychological Science and of the Psychonomic Society and Sequoyah Fellow of the American Indian Science and Engineering Society.

Kelly Jakubowski is Assistant Professor of Music Psychology at Durham University, UK, where she serves as the co-director of the Music & Science Lab (musicscience.net). She currently holds a Leverhulme Early Career Fellowship titled "Prevalence, features, and retrieval of music-evoked autobiographical memories." This project employs multiple methodologies for collecting a large and diverse dataset of lifetime memories triggered by listening to music, with the aims of gaining a more systematic understanding of the conditions under which these memories occur and expanding theoretical accounts of the interactions among music, memory, and emotions. Other research interests include musical imagery, involuntary memory, timing and synchronization in music performance, and cross-cultural music perception.

Jin Hyun Kim, PhD, having worked at several transdisciplinary research centres on media and communication, (embodied) music cognition, languages of emotion, and cognitive sciences and neurosciences, has been developing a programme of philosophically grounded and empirically oriented basic music research as Director of the Lab of Music Technology, Aisthesis, and Interaction (TAIM) at Humboldt-Universität zu Berlin. Currently, she is heading the research project "Social interaction through sound feedback—Sentire" funded by the German Federal Ministry of Education and Research.

Mahiko Konishi is a psychology researcher and behavioural scientist based in Paris, France. His interests focus on metacognition and mind-wandering, but he likes to tackle different topics in experimental psychology and has also worked on false memories, stress, attention, and multitasking. Mahiko recently worked at the Centre d'Économie de la Sorbonne and at the École Normale Supérieure in Paris, France; he received his PhD degree from the University of

York (UK), studying the neural bases of mind-wandering under the supervision of Dr Jonathan Smallwood.

Mats B. Küssner is Lecturer in the Department of Musicology and Media Studies at Humboldt-Universität zu Berlin and Visiting Research Fellow in the Department of Psychology at Goldsmiths, University of London. His research focuses on multimodal perception and mental imagery of music, emotional responses to music, and performance science. Before moving to Berlin, he was Peter Sowerby Research Associate in Performance Science at the Royal College of Music. His work has been published in *Psychology of Music*, *Music Perception*, *Musicae Scientiae, Empirical Musicology Review*, *Music & Science*, *Psychomusicology*, *Frontiers in Psychology*, *NeuroImage*, and *PLOS ONE*.

Bence Nanay is BOF Research Professor at the Centre for Philosophical Psychology at the University of Antwerp and Senior Research Associate at Peterhouse, Cambridge University. He is the author of three books (*Between Perception and Action*, Oxford University Press, 2013; *Aesthetics as Philosophy of Perception*, Oxford University Press, 2016; *Aesthetics: A Very Short Introduction*, Oxford University Press), with six more under contract (four for Oxford, one for Routledge, and one for W. W. Norton) and more than 140 peer-reviewed articles. His work is supported by a number of high-profile grants including an ERC Consolidator Grant.

Jonathan M. Oakman is Associate Professor in the Psychology Department at the University of Waterloo in Ontario, Canada. He is also a practicing clinical psychologist. Unlike most of the other contributors to this volume, he has no discernible musical talent and would struggle to play the kazoo.

André Ofner is a PhD student at the Artificial Intelligence Lab at Otto-von-Guericke University in Magdeburg. He investigates computational models of cognition at the intersection of neuroscience and artificial intelligence with a focus on predictive processing and music perception. He holds a bachelor's degree in bioinformatics from the Technical University of Munich, Germany and a master's degree in cognitive science from the University of Vienna, Austria.

Konstantina Orlandatou studied composition, music theory, piano, and accordion in Athens, multimedia composition (MA degree) at the University of Music and Drama in Hamburg and completed her doctoral studies (PhD degree) in systematic musicology at the University of Hamburg (Dissertation: "Synaesthetic and intermodal audio-visual perception: an experimental research"). As an active composer and musicologist, her research focuses mainly on audio-visual interaction and perception. Since 2018, she has been leading the project "Moving Sound Pictures" at the University of Music and Drama in Hamburg, where she uses virtual reality and its technologies as a new form of interactive art mediation between music and visual arts.

Dr Graziana Presicce recently completed a scholarship-funded PhD degree in music performance at the University of Hull. Her doctoral research investigated music listeners' experiences of visual imagery and engagement in response to classical piano music. Besides her active interests in music and visual imagery, Graziana recently collaborated as Research Assistant in the STROKESTRA research project: an RPO/SEMPRE-funded investigation exploring the implementation and impact of a post-stroke orchestral rehabilitation programme. As a classical pianist, Graziana frequently engages in recitals, both as a soloist and as an accompanist and chamber musician.

Rebecca S. Schaefer has a background in clinical neuropsychology, music cognition, and cognitive neuroscience. She started researching music imagery in the context of brain–computer interface research, and has since investigated music imagery with several brain-imaging methods and behavioural- and questionnaire-based research. Currently in the Neuropsychology Department of Leiden University, The Netherlands, she focuses on health applications of musical interactions, specifically involving music imagery, perception, and moving to musical rhythm, using neuroimaging, behavioural, and cognitive measures.

Sebastian Stober is Professor for Artificial Intelligence at the Otto von Guericke University Magdeburg. He studied computer science with a focus on intelligent systems and a minor in mathematics in Magdeburg until 2005 and received his PhD degree on the topic of adaptive methods for user-centred organization of music collections in 2011. From 2013 to 2015, he was a postdoctoral fellow at the Brain and Mind Institute in London, Ontario, where he pioneered deep learning techniques for studying brain activity during music perception and imagination. Afterwards, he was the head of the Machine Learning in Cognitive Science Lab at the University of Potsdam, before returning to Magdeburg in 2018.

Liila Taruffi is Lecturer in Music Psychology at Durham University, UK. She has an interdisciplinary background in psychology, neuroscience, and aesthetics. Her current research is centred on music's capability to influence a range of internally oriented mental phenomena such as mind-wandering and visual mental imagery. Ultimately, her research is driven by the goal of practical applications of music to help maximize the beneficial outcome of mind-wandering and visual imagery and overcome their detrimental aspects. Other research interests range over empathic responses to music, creativity, music and well-being, and alexithymia. Liila uses the tools of affective and cognitive psychology and neuroimaging techniques such as fMRI. She also employs ecologically valid methodologies, such as mobile experience sampling in daily life, psychophysiology, and analysis of qualitative data.

William Forde Thompson is Director of the Music Lab at Bond University, Gold Coast, Australia and the Music Sound and Performance Lab at Macquarie University, Sydney, Australia. He is the author of "*Music, Thought and Feeling: Understanding the Psychology of Music*" (2nd edition), the Editor of *The Science and Psychology of Music: From Beethoven in the Office to Beyoncé at the Gym* (2021, with co-Editor Kirk N. Olsen), and the Editor of the *Sage Encyclopedia of Music in the Social and Behavioral Sciences* (2014). He has served as President of the *Australian Music Psychology Society, The North American Society for Music Perception and Cognition*, and the *Asia Pacific Society for the Cognitive Sciences of Music*.

Thijs Vroegh has a background in economics and musicology and worked among others as a statistical programmer in the domain of clinical studies. At the moment, he is working as a postdoctoral researcher in the Music Department of the Max Planck Institute for Empirical Aesthetics in Frankfurt am Main. In 2018, he successfully finished his dissertation (summa cum laude) on the aesthetic aftereffects of being absorbed in music. His current research is centred on the phenomenological correspondences between intense music listening and hypnosis, combining structural equation modelling, psychological network analysis, and causal discovery algorithms.

Silvia van Zyl obtained a Performance Diploma degree in music (with distinction) in 2016, a BMus Honours degree (*cum laude*) in 2017, and an MMus degree (*cum laude*) degree at the University of Cape Town in 2020. She is currently pursuing a PhD (Musicology) at the University of Cape Town with research focusing on mental imagery in visually impaired pianists. In 2018, she was the winner of the continent-wide essay competition in the *Journal of the Musical Arts in Africa* and has been awarded numerous scholarships throughout her undergraduate and postgraduate studies.

Acknowledgements

We owe a great debt to a number of people who supported us along the way and made this volume possible. First and foremost, we thank all our contributors for sharing their interest in research on music and mental imagery and for producing 24 thought-provoking chapters which allow the reader to explore this topic from a range of different perspectives. We are very grateful to our expert reviewers for their time and constructive feedback on individual chapters of this volume. They are (in alphabetical order): Mihailo Antović, Krystian Barzykowski, Amy Belfi, Charles-Étienne Benoit, Laura Bishop, Olivier Brabant, Marilyn Clark, Terry Clark, Eric Clarke, Tom Cochrane, Caroline Curwen, Mine Doğantan-Dack, Kathryn Emerson, Jörg Fachner, Georgia A. Floridou, Takako Fujioka, Emma Greenspon, Timothy L. Hubbard, Kelly Jakubowski, Petr Janata, Peter Keller, Mariusz Kozak, Dan Leech-Wilkinson, Clemens Maidhof, Corinna Martarelli, Aaron Mitchel, Daniel Müllensiefen, Bence Nanay, Diana Omigie, Marcus Pearce, Lara Pearson, Timothy Pruitt, Frances Ragni, Adriana Renero, Mark Reybrouck, Nicole Saintilan, Rebecca S. Schaefer, Bill Thompson, Renee Timmers, Michelle Ulor, Manila Vannucci, Floris van Vugt, Edward Vessel, Johan Willander, and two anonymous reviewers. We would also like to express our gratitude to Andrea Halpern and Elizabeth Margulis for action-editing the chapters we (co-)authored ourselves. Special thanks go to a group of advanced MA students in musicology at Humboldt-Universität zu Berlin who provided critical comments on a draft of the whole volume. We are particularly grateful to Olivia Geibel for taking care of style and formatting of the manuscript in the weeks leading up to submission—without her help, we'd still be digging through the APA publication manual. At Routledge, we thank Heidi Bishop and Kaushikee Sharma for their expert advice and patience with our queries. We also gratefully acknowledge the generous financial support by the Chair of Transcultural Musicology and Historical Anthropology of Music at Humboldt-Universität zu Berlin and by the Department of Music at Durham University. Finally, we are enormously indebted to our colleagues, friends, and families for their encouragement during any stage of the writing and editing process.

Foreword

Mental imagery (MI) is a rich internal experience, which can comprise several modalities, most commonly visual, motor, and auditory domains. It can be experienced spontaneously or deliberately, can be neutral or emotional, and can be helpful or confounding to an ongoing task. All of these characteristics and more are often experienced as a reaction to hearing, performing, remembering, or composing music, the topic of the current volume.

Although philosophers and others have described MI in historical and current eras, the rigorous scientific study of MI has a more recent history. Because it is an internal experience without obvious external manifestations, experimental psychologists needed to first develop ways of documenting even the existence and then the characteristics of MI. How do I know what you are doing when you say you are "hearing" a song in your head, or visualizing a scene to music? Imagery specifically relating to music is also a challenge to capture and study, because it sometimes involves emotions. And another complexity is that sometimes MI is under voluntary control such as deliberate recall of a favourite song, but other times seems to arrive unbidden, either spontaneously such as during mind-wandering or in response to a cue (Floridou in the current volume).

My interest in studying MI, particularly auditory imagery for music, came from my early and profound interest in memory (MI is often considered variety of memory). One of my teachers in graduate school, Roger Shepard, was among the first experimental psychologists to develop techniques to externalize the experience of MI, by for instance measuring how long it took people to mentally rotate complex figures. This inspired me and others to devise ways of objectively measuring similar experiences when people imagine music, for instance tapping along to an imagined melody, or mentally transposing or reversing a melody (Gelding, Day, and Thompson). Although people are not necessarily fully aware of their MI, some aspects are very conscious and thus verbalizable. This has led to interesting investigations using self-report, either in the form of questionnaires or via qualitative analyses of descriptions of their MI, including in different states of consciousness such as musical daydreaming (Konishi; Herbert; and Fachner). I find it very encouraging that more objective and more subjective behavioural data are now viewed as complementary rather than antagonistic as has sometimes been the case, and that increasingly rigorous ways of dealing with coding of qualitative data and norming of questionnaires have been applied to studies of musical MI, such as coding the content of music-evoked autobiographical memories (Jakubowski) or indexing the vividness or clarity of musical imagery (Hubbard). A third methodological perspective of increasing importance is the neuroscience of MI, which is often combined with the behavioural methods just mentioned to form a more complete picture of how imagery uses some of the neural networks that are used in perception of music or performing, but it also reflects the unique aspects of that mental simulation (Belfi; Ofner and Stober).

Musical imagery during performance is an area that I've been studying recently. Consider informal singing. Even casual singers require motor imagery for pitch matching and breath control, as they cannot see any of the body parts that produce the sound. In more formal situations, choral or solo singers employ auditory imagery from a visual cue as they read notation, motor imagery as they anticipate reacting to the conductor's next move, or they may visualize sheet music. For instrumentalists, hands and other visible body parts need to act in exquisite sequence (percussionists even have to transit from one instrument to another) and thus motor and visual imagery are likely very important (discussed by Black; Presicce; and Herbst and van Zyl). It would be interesting to consider MI in the context of musical theatre, where music and movement need to be coordinated in acting and even dancing.

I mentioned earlier that MI can also be associated with emotion. In one of my studies, I was struck by the extreme similarity of real-time judgements of conveyed emotion, in heard and imagined music, resulting nearly identical timing profiles (Lucas et al., 2010). Music intended for an audience of course needs to convey emotion, which may require a mental simulation by the performer, because that may not be the emotion the performer is feeling. For instance, as a singer, my MI might be needed to properly prepare for that difficult high but soft note, by modelling motor commands, and anticipating the conductor's gesture. But I also use MI to simulate and then convey an emotion, if that is what the music intends. MI can also induce emotion: People can play mental songlists to alter mood, including in difficult conditions such as isolation or during extreme stress. This aspect of musical MI can also be useful in therapeutic situations (Dukić; Finch and Oakman; Schaefer).

Particular episodes of MI need not be crystallized experiences but can be part of a dynamic system such as in composing. I recently came across a compendium of letters/diaries from nearly 100 years ago by Marie Agnew (Agnew, 1922), in which the author presented extracts of letters and diaries by five great composers that reflect what we would call musical MI. Mozart's comments on his imagery are widely known, but I was struck by the clarity (and utility) of auditory imagery reported by Schumann: "When you begin to compose, do it all with your brain. Do not try the piece at the instrument until it is finished." And the complexity of what Tchaikovsky imagined: "Thus I thought out the scherzo of our symphony—at the moment of its composition, exactly as you heard it. It is inconceivable except as a pizzicato." These quotes mostly reflect auditory imagery, but Wagner is quoted as integrating visual and auditory imagery:

> There rose up soon before my mind a whole world of figures . . . that when I saw them clearly before me and heard their voices in my heard, I could not account for the almost tangible familiarity and assurance in their demeanor.

This multimodality of imagery with respect to music is of increasing interest in this field, as reflected in chapters by Taruffi and Küssner; Godøy; Kim; and Nanay.

One aspect of this field that I'd like to encourage for the future is a closer examination of individual differences in frequency and characteristics of MI for music. We may not be as aware of interindividual differences compared to other skills, because of their covert nature: it may not even occur to me that my mental image of Happy Birthday might differ from yours. Many of the approaches described in this volume are very suitable for documenting differences in vividness and accuracy, and the ways musical MI may be used by different people, in different situations. Also valuable would be studying developmental trajectories in musical MI, from childhood to older age. These perspectives will add to our understanding of the connections between particular experiences of musical MI, and other perceptual, cognitive, creative, aesthetic, and therapeutic consequences of this rich experience.

Andrea R. Halpern

References

Agnew, M. (1922). The auditory imagery of great composers. *Psychological Monographs*, *31*, 279–287. https://doi.org/10.1037/h0093171

Lucas, B. J., Schubert, E., & Halpern, A. R. (2010). Perception of emotion in sounded and imagined music. *Music Perception*, *27*, 399–412. https://doi.org/10.1525/mp.2010.27.5.399

Preface

The idea of this volume emerged after two KOSMOS[1] meetings which were held at Humboldt-Universität zu Berlin in 2017 and 2018. Inspired by Juslin and Västfjäll's (2008) framework on how music induces emotion in the listener, our initial focus was on music, emotion, and visual mental imagery. Juslin and Västfjäll postulated that visual imagery[2] (i.e., images in the mind's eye) is one of several mechanisms that evoke an emotional response in the listener, but few attempts had been made to investigate this mechanism (Juslin et al., 2008; Vuoskoski & Eerola, 2015). To tackle this gap of knowledge, we organized the KOSMOS Dialogue "Music, Emotion, and Visual Imagery" at Humboldt-Universität zu Berlin in June 2017. This workshop brought together psychologists, musicologists, music therapists, and other music scholars, paved the way for future collaborations, and resulted in a special issue on music, emotion, and visual imagery (Küssner et al., 2019). However, this meeting also revealed some of the methodological issues (e.g., how to reliably measure visual imagery) and blurry boundaries between (visual) mental imagery and other mental phenomena such as mind-wandering and autobiographical memory. Furthermore, we realized that the gap of knowledge concerns not only the mechanism of how visual imagery induces emotion during music listening but also what visual imagery in a musical context actually consists of (and for whom and why). There is a substantial amount of literature on (visual) mental imagery in the cognitive (neuro-)sciences (Kosslyn et al., 2001; Pearson, 2019; Pylyshyn, 2002; Thomas, 1999) and philosophy (Gregory, 2010; Nanay, 2015), but research on visual imagery in musical contexts was notably absent, with the exception of music education research concerned with mental imagery skills as a way to enhance performance practice, for example, by visualizing musical scores or body movements (Clark et al., 2011). These observations led us to organize the KOSMOS Workshop "Mind Wandering and Visual Mental Imagery in Music"[3] in May 2018, with a broader focus on visual imagery and related states of consciousness associated with musical activities. Specialists presented their work on mind-wandering, motor rehabilitation, musical daydreaming, distraction during musical performances, and music therapy to name but a few topics, and there was a sense that all of these works were related, but different terms, methods, and theories rendered it difficult to pinpoint how these eclectic approaches have the potential to inform one another. The impression after the workshop was that we need to consider all forms of mental imagery to be able to accommodate the variability of these research activities and start developing a more coherent, interdisciplinary research framework of music and mental imagery. This volume, with many contributions by scholars who presented their work at the aforementioned meetings, is our first attempt. Its main goal is to bring together—and critically discuss—theoretical and empirical work on mental imagery associated with, or related to, music, and develop potential avenues for future collaborative work where each participating discipline gives and gains something.

Notes

1 The KOSMOS programme consisting of three different formats (dialogue, workshop, and summer university) was developed to create sustainable international networks in research and teaching and was funded through Future Concept resources of Humboldt-Universität zu Berlin through the Excellence Initiative of the German Federal Government and its Federal States.
2 We use the terms "visual mental imagery" and "visual imagery" synonymously.
3 Further materials on both KOSMOS meetings can be found online at https://sites.google.com/view/emuvis/kosmos-workshops

References

Clark, T., Williamon, A., & Aksentijevic, A. (2011). Musical imagery and imagination: The function, measurement, and application of imagery skills for performance. In D. J. Hargreaves, D. Miell, & R. A. R. MacDonald (Eds.), *Musical imaginations: multidisciplinary perspectives on creativity, performance and perception* (pp. 351–365). Oxford: Oxford University Press.

Gregory, D. (2010). Imagery, the imagination and experience. *The Philosophical Quarterly*, *60*(241), 735–753. https://doi.org/10.1111/j.1467-9213.2009.644.x

Juslin, P. N., Liljeström, S., Västfjäll, D., Barradas, G., & Silva, A. (2008). An experience sampling study of emotional reactions to music: Listener, music, and situation. *Emotion*, *8*(5), 668–683.

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, *31*(5), 559–575.

Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience*, *2*(9), 635–642. https://doi.org/10.1038/35090055

Küssner, M. B., Eerola, T., & Fujioka, T. (2019). Music, emotion, and visual imagery: Where are we now? *Psychomusicology: Music, Mind, and Brain*, *29*(2–3), 59–61. https://doi.org/10.1037/pmu0000245

Nanay, B. (2015). Perceptual content and the content of mental imagery. *Philosophical Studies*, *172*(7), 1723–1736. https://doi.org/10.1007/s11098-014-0392-y

Pearson, J. (2019). The human imagination: The cognitive neuroscience of visual mental imagery. *Nature Reviews Neuroscience*, *20*(10), 624–634. https://doi.org/10.1038/s41583-019-0202-9

Pylyshyn, Z. W. (2002). Mental imagery: In search of a theory. *Behavioral and Brain Sciences*, *25*(2), 157–182. https://doi.org/10.1017/S0140525X02000043

Thomas, N. J. T. (1999). Are theories of imagery theories of imagination? An active perception approach to conscious mental content. *Cognitive Science*, *23*(2), 207–245. https://doi.org/10.1207/s15516709cog2302_3

Vuoskoski, J. K., & Eerola, T. (2015). Extramusical information contributes to emotions induced by music. *Psychology of Music*, *43*(2), 262–274. https://doi.org/10.1177/0305735613502373

Introduction and Overview

Mats B. Küssner, Liila Taruffi, and Georgia A. Floridou

Most of us would have experienced a piece of music stuck in our head (a so-called "earworm"; Floridou et al., 2015; Liikkanen & Jakubowski, 2020), playing over and over in our mind for days if not weeks. Perhaps, it is a song that we first heard at a concert some days earlier and that our mind seems to automatically and spontaneously recreate for us—whether we like it or not. Our capacity to "hear" music in the absence of the corresponding external sound waves is a prime example of auditory, or indeed, musical imagery, and belongs to the broader category of mental imagery. In most general terms, *mental imagery* is a sensory-like experience in a modality in the absence of an actual corresponding sensory input in that modality. This definition is in line with others from the cognitive sciences such as the one provided by Kosslyn et al. (1995, p. 1335):

> [v]isual mental imagery is 'seeing' in the absence of the appropriate immediate sensory input, auditory mental imagery is 'hearing' in the absence of the immediate sensory input, and so on. Imagery is distinct from perception, which is the registration of physically present stimuli.

However, scholars from other (applied) disciplines have different connotations with, and perspectives on, the term *mental imagery*. The following definition from the realm of music education stresses the practical function of mental imagery in a typical rehearsal context: mental imagery is characterized as

> cognitive or imaginary rehearsal of a physical skill without overt muscular movement. The basic idea is that the senses—predominantly aural, visual, and kinesthetic for the musician—should be used to create or recreate an experience that is similar to a given physical event.
>
> (Connolly & Williamon, 2004, p. 224)

To carve out the subtle (or not-so-subtle) differences of scholars' perspectives on mental imagery, we asked each contributor of this volume to provide their own definitions—resulting in 24 nuanced and multifaceted versions from fields such as psychology, philosophy, musicology, music therapy, and music education.

DOI: 10.4324/9780429330070-1

Mental Imagery: Ordinary or Originary?

While some forms of mental imagery, as in the case of earworms, may be based on past experiences and mechanisms of memory retrieval (Liptak et al., 2022) and maintenance (Killingly et al., 2021), our mind is capable of much more: humans—especially trained musicians—can voluntarily imagine a melody in various timbres and turn, for example, a piano tune into a full-blown orchestra version. Or we can create new sounds in our mind's ear that have never been instantiated in physical sound waves, which is a typical mental practice of both composers and musical performers. However, some might argue that all personal musical imagery must be based on, or related to, one's previous listening experiences. This refers to the broader philosophical question of "whether imagination is autonomous and originary or rather is derivative of other psychological processes and hence ordinary" (Waller et al., 2012, p. 298) and whether creative imagination can be measured with scientific methods at all (Thomas, 1999). The separation of mental imagery based on memory and creative processes has a long tradition in the literature and comes with various terms and definitions. For instance, already James (1890) distinguished between *reproductive* and *productive imagination*; more recent accounts use the terms *memory imagery* and *imagination imagery* (Gracyk, 2019). As further discussed in Waller et al. (2012), contemporary scientific endeavours treat mental imagery mostly as "memory images," that is, as based on previous experiences, whereas imaginative, originary accounts had been ignored. The chapters offered in this volume do not support such a view. Many accounts and research reports provided throughout this volume testify to the power of music to create genuinely new mental images that are not simply reinstantiations of previous external sensory input in the mind's eye or ear. Music may, in that sense, reveal itself as one key to unlock the gate to originary mental processes that engage one or several modalities. Indeed, the fact that mental imagery in musical contexts occurs not only in the auditory but also in visual, kinaesthetic, or even tactile modalities implies that a multimodal perspective is needed for a comprehensive understanding of mental imagery (Nanay, 2018).

Multimodality

The importance of multimodality has been discussed by Godøy and Jørgensen (2001) some 20 years ago and is taken up again in this volume. Conceived by various music scholars in the late 1990s, Godøy and Jørgensen's volume *Musical Imagery* aimed to (at least partially) rectify cognitive scientists' bias towards visual imagery in their pursuit of shedding light on the mechanisms underlying mental imagery. However, as the editors of that volume soon noted while organizing the structure of the book, a narrow focus on auditory mental imagery could not possibly account for the complex mental processes involved in music listening and making. Indeed, when we listen to music, we may conjure up visual images of, for instance, performing musicians or of natural landscapes in our mind's eye (visual mental imagery); feel a sense of movement that corresponds to the temporal unfolding of a piece (kinaesthetic mental imagery); or may even imagine a distinct smell (olfactory mental imagery) that we associate with the emotional character of the music or that has been triggered by an episodic memory. Although a recent study (Wilain et al., 2021) has shown that visual and kinaesthetic mental imagery are the two modalities most often involved during music listening, other modalities such as olfactory, gustatory, and tactile do occur as well. When mental images in two or more modalities are formed and experienced simultaneously or in succession, this is what we refer to as *multimodal mental imagery*. For instance, when we hear *and* see a musical performance in our mind, this is multimodal mental imagery—regardless of whether an external stimulus (such as music) is present or not. This definition differs from Nanay's definition (this volume) and highlights why it is so important to define terms and concepts precisely. Nanay argues that music-evoked visual mental imagery *in itself* is a case of multimodal mental imagery, because two modalities are involved: auditory (perception) and visual (imagery). If Nanay (2021) is right that some forms of mental imagery are unconscious, it will be challenging to ascertain which modalities are involved during music listening. What seems clear, though, is that all these mental images, stemming from various modalities, are an integral part of musical experience, highlighting that multimodality is a central feature of music-related mental imagery.

From Basic to Applied Research in Mental Imagery

Although Küssner and Orlandatou (this volume) briefly touch on the history of mental imagery research and Taruffi and Küssner (this volume) describe the basic points of the "imagery debate" that captured cognitive science for many years (Kosslyn et al., 2006), most chapters are concerned with the recent academic discourse. For an excellent summary of the beginnings of empirical research on mental imagery at the end of the 19th century—when Carl Stumpf and his contemporaries paved the way for much of the experimental music research as we know it today—and all the philosophical writings on mental imagery that preceded it, we refer the reader to Schneider and Godøy (2001). Suffice it to say that, in his essay "On Gestalt Qualities" in 1890, Christian von Ehrenfels, the philosopher and early protagonist of Gestalt psychology, already emphasized the importance of multimodal mental imagery for musical works. He stressed that, when thinking about an orchestral piece, one should imagine a scene with the entire orchestra, the concert setting, the lighting, the movements of the conductor and musicians, and so forth. Some of the questions and issues discussed in this volume therefore have a long history and engaged generations of scholars interested in the fundamental mechanisms of how music plays (in) our mind and their relation to cognitive and affective processes. It is with this historical awareness that we present the current volume and hope that a fresh approach to the topic with new interconnections between related psychological phenomena and a glimpse into areas of applied, music-related mental imagery research will provide a useful resource for students, scholars, and music practitioners. Testimony to the applicability of music and mental imagery research already comes from translational research developed in this area. For instance, recent years have seen a rise of studies on auditory imagery (Halpern & Overy, 2019; Schaefer, 2017; Watanabe et al., 2020) including applications of how imagined sound and music can aid in pedagogy, clinical contexts, and rehabilitation. While each individual chapter is self-contained and can be enjoyed on its own, we expect the number of revelatory insights into music-related mental imagery to correlate positively with the number of chapters read.

Part I: Modalities of Mental Imagery

The first part of the present volume provides an overview of the most commonly experienced types of modalities of mental imagery in musical contexts (see Wilain et al., 2021), namely auditory, visual, and kinaesthetic. As we perceive music holistically, a division into separate sense modalities is somewhat artificial, but it helps to delineate basic concepts and theories. It also reflects the different research strands that are usually concerned with one specific type of mental imagery. The last chapter of Part I provides one account of how multimodal mental imagery may work in musical contexts.

Introducing the reader to the most commonly studied form of mental imagery in relation to music, that is, musical imagery, Floridou develops, in Chapter 1, a rich conceptual framework of its multifaceted forms ranging from voluntary musical imagery during mental rehearsal over anticipatory imagery during music listening to spontaneously occurring musical mind-pops and earworms. She critically reviews literature on musical imagery occurring before, during, and after musical activities and shows how its various forms are intricately linked with everyday cognitive processes such as mind-wandering, memory, or future thinking. Above all, her analysis reveals that musical imagery is an umbrella term for many different concepts that musically trained and untrained individuals alike regularly encounter in everyday life.

In Chapter 2, Taruffi and Küssner address visual mental imagery (VMI) in various musical contexts, ranging from shamanic rituals to music therapy, and emphasize the individuality of VMI experiences during music listening. Music-related VMI can be associated with a number of related cognitive phenomena such as mind-wandering, absorption, or trance, which will be dealt with in depth in Part III of this volume. The first main section of this chapter is devoted to the format of mental representations underlying VMI, as Taruffi and Küssner re-visit the "imagery debate" (Kosslyn et al., 2006). The basic question of the imagery debate is whether mental representations are depictive or propositional, that is, whether information is encoded image-like or with abstract, linguistic symbols. There is now a substantial amount of neuroscientific evidence in favour of the depictive account, demonstrating that topographic areas of the visual cortex are activated during both perception and imagery. The role of affect in music-related VMI is the focus of the second main section. Taruffi and Küssner introduce Juslin and Västfjäll's framework of emotion induction during music listening

(Juslin & Västfjäll, 2008) and report several recent empirical studies which have begun to shed light on the relationship among VMI, music, and emotion. The last section is concerned with various clinical applications—from music therapy to chemotherapy—in which VMI already plays a prominent role and which could be further enhanced by music-based interventions. As well as highlighting the manifold forms and functions of music-related VMI, the chapter anticipates some of the topics that are discussed in more detail in Parts III and IV of this volume.

In Chapter 3, Godøy illuminates musical imagery—similar to Floridou in Chapter 1—but with a focus on the connection between actions and sounds, building on his extensive work on sound-motion objects. *Sound-motion objects* have a duration of 0.3–3 seconds and "are perceived and conceived holistically as coherent units, and notably so, as units combining sensations of sound and body motion" (Godøy, 2019, p. 161). Heavily influenced by the theoretical work of Husserl and Schaeffer, Godøy argues that musical imagery—as well as music perception in general—is shaped by constraints of our motor system and that mental imagery of sound-producing actions can trigger mental imagery of musical sounds. His central theoretical framework is *intermittent motor control*, that is, the idea that actions are planned, executed, and monitored in chunks of variable length. These can then, in turn, be found in our perception of musical shapes which are similarly organized in sound-motion objects at different timescales. Employing a combination of music theory, phenomenology, and cognitive science, Godøy situates volitional musical imagery within broader processes of musical creativity and provides robust empirical evidence for his thesis that sound-motion objects are central ingredients of our musical listening and mental imagery experiences.

Drawing on the work of Godøy, Cox, Lipps, and Stern among others, Kim develops, in Chapter 4, an intricate account of kinaesthetic musical imagery that is distinct from commonly encountered concepts such as kinaesthetic imagery and motor imagery, which are often used synonymously in the music literature and refer to first- or third-person perspectives of human body movements. Disentangling those terms, she puts her finger on various inconsistencies in the literature and delineates her own concept that is informed by historical discourses—for instance, Lipps' discussion of the term "Einfühlung" (i.e., empathy)—as well as her own empirical work using micro-phenomenological interview techniques with musicians to carve out second-person perspectives of music cognition. Because kinaesthesia and kinaesthetic imagery have so many different connotations in the literature, Kim feels obliged to outline first what her concept of kinaesthetic imagery does *not* involve. For instance, it is not the simulation of motor actions that are necessary for sound-producing gestures and it is also not the conscious imagination of (imitated) motor actions. Rather, according to Kim, "*kinaesthetic musical imagery* can be understood as a category of mental imagery that could be best characterized as the (quasi-)perceptual conscious experience of dynamic self-movement that mental representations of musical dynamic properties give rise to." She contextualizes her approach within theories of consciousness and draws connections to Stern's psychodynamic forms of vitality, offering a new perspective on a concept—kinaesthetic imagery—which, on close inspection, proves to be rather thorny and theoretically challenging.

In Chapter 5, Nanay draws on philosophy, psychology, and neuroscience to provide his account of *multimodal mental imagery* in musical contexts. While mental imagery may be triggered top-down (e.g., by voluntarily imagining a national anthem), multimodal mental imagery, according to Nanay, is triggered laterally by sensory input from a modality (e.g., vision) that is different from the relevant modality (e.g., audition) in which mental imagery occurs. Nanay outlines some general features of mental imagery: it can be voluntary or involuntary, conscious or unconscious, and the location of its object can be internal (e.g., before the mind's eye) or external (e.g., on the desk or the wall in front of us). He discusses the role of (musical) expectations and how some belong to the category of *auditory temporal mental imagery*, that is, perceptual processing that comes either too late or too early, but does not correspond to the relevant sensory input. Nanay reports some neuroscientific evidence showing that our brain is able to process perceptual information of highly familiar stimuli before the corresponding sensory input has occurred. A convincing musical example is the installation *Earth-Moon-Earth (Moonlight Sonata Reflected From The Surface of The Moon)*, a piece of sound art where our expectations fill in missing notes of the Moonlight Sonata. In the remainder of the chapter, Nanay draws on various illustrative examples of music and dance performances to argue that multimodal perception of music is the norm and can have—enabled by multimodal mental imagery—lasting effects on our musical experience. For instance, he shows how dancers' gestures can trick us into hearing altered time signatures, and how this mode of listening to the musical metre may persist—due to (unconscious) visual mental imagery of those gestures—when being presented with an audio-only version of that performance some weeks later.

Part II: Measurement

The second part is concerned with the tricky question of how to measure and analyse experiences of music-related mental imagery. Experts in their respective fields reveal advantages and disadvantages of state-of-the-art tools and methods ranging from verbal self-report to neuroimaging approaches and deep neural networks. All measurements are discussed with respect to relevant empirical findings from research on music-related mental imagery. The last chapter of this part encompasses a broader, cross-cultural perspective and emphasizes some important ethical questions (e.g., power relations) relevant not only to field work in remote areas of the world but also to the daily work in music psychological laboratories.

In Chapter 6, Gelding, Day, and Thompson introduce subjective and behavioural measures of music-evoked visual imagery and auditory imagery for music, highlighting a number of imagery-related dimensions that researchers usually assess with self-reports. Dimensions of music-evoked visual imagery include, for instance, prevalence, nature, content, quality, vividness, intensity, timing, and duration. Gelding and colleagues focus on temporal dynamics of visual imagery and report one of their own studies (Day & Thompson, 2019) that used chronometric measures (i.e., reaction times) to investigate when people experience emotional responses and visual imagery during music listening. Chronometry offers a promising way to study the causal relationship between visual imagery and felt emotions, but a variety of methods (e.g., supressing visual imagery, see Hashim et al., 2020) are needed before more solid conclusions about causality can be drawn. The second part of their chapter is concerned with auditory imagery for music, both during and after listening. The authors emphasize the need to develop more robust, adaptable behavioural methods that take into account interindividual differences and ensure that people actually use musical imagery to solve a given task. Two novel methods developed by the authors—the Pitch Imagery Arrow Task (PIAT) and the Rhythm Imagery Task (RIT)—are described to address the aforementioned issues. Interestingly, Gelding and colleagues' use of the term *anticipatory imagery* (i.e., auditory imagery during music listening) seems to largely overlap with Nanay's concept of *auditory temporal mental imagery*, as both deal with the brain's anticipated continuation of musical stimuli. The authors close their chapter with a plea to make use of more neuroscientific methods (see Belfi, this volume) to corroborate subjective and behavioural measures.

In Chapter 7, Hubbard provides a systematic overview of, and critical reflection on, self-report measures of mental imagery in musical contexts, not only emphasizing some of the points made by Gelding et al. (this volume) but also introducing a number of novel aspects. Under the term *self-report measures*, Hubbard subsumes both verbal questionnaire data and non-verbal behavioural responses. He discusses in detail the advantages and disadvantages of verbal self-report and pays attention to disadvantages that are specific to music and imagery research such as the insensitivity of questionnaires to transient episodes of visual imagery, the blurring of imagery and non-imagery information, and the issue of participants reporting their interpretation of visual imagery (rather than the visual images themselves). Notable behavioural measures include key presses (judgements and selections), drawing lines, adjusting metronomes, and finger tapping. In the second part of his chapter, Hubbard outlines musical and non-musical dimensions of musical imagery, where *musical imagery* denotes a broad spectrum of different types (usually visual and auditory) of mental imagery. Musical dimensions include pitch, timbre, contour, tempo, loudness, expressive timing, tonality, and lyrical content; non-musical dimensions encompass vividness, clarity, valence, ability to control or transform (auditory) images, and voluntariness. He emphasizes that self-reports not only complement neuroscientific methods and form an integral part of recent methodological developments (e.g., neuro- or heterophenomenology) but also allow researchers to track emergent properties of mental imagery.

Belfi focuses on neuroscientific measures in Chapter 8, aiming to provide practical guidance in choosing the appropriate method for identifying neural correlates of music-related mental imagery. The first approach discussed is lesion studies with patients suffering from focal brain damage due to strokes, removal of brain tissue, or infections. For instance, one such neuropsychological study revealed that musical hallucinations—a form of musical imagery where people imagine music coming from the external world—are associated with damage to the (left) temporal lobes. A method more commonly used in mental imagery research is functional magnetic resonance imaging (fMRI). A typical question that can be addressed in fMRI studies is whether perception and mental imagery of music (or other stimuli) rely on the same neural structures. Findings have revealed that there is substantial overlap, but mental imagery does not necessarily activate primary sensory regions such as the primary auditory cortex for musical imagery. Belfi also discusses electroencephalography (EEG), which has some advantages in comparison with fMRI (cheaper, silent, mobile devices available), and reports a couple of findings from musical imagery studies. For instance, Schaefer et al. (2011) showed that musical imagery compared to perception gives rise to more activation in the alpha frequency range (8–12 Hz) over occipito-parietal areas, indicating that musical imagery involves working memory processes. The last method covered is magnetoencephalography (MEG). Belfi discusses one MEG study (Gelding et al., 2019) utilizing the Pitch Imagery Arrow Task, which revealed that beta activity (13–30 Hz) in auditory and sensorimotor areas is associated with mental imagery for musical pitch. To close her chapter, she cautions against using neuroscientific measures without accompanying self-reports or behavioural methods, thereby echoing arguments made by Hubbard (this volume) and Gelding et al. (this volume).

In Chapter 9, Ofner and Stober delve deeper into neuroimaging data and provide an overview of the latest machine learning techniques (i.e., extracting knowledge from raw data) to make sense of EEG and fMRI data collected during music listening. Machine learning algorithms are used to find patterns and trends in raw data that are not visible to the naked eye, with the aim to understand the data better or to make predictions based on the detected patterns. In the realm of music research, one of the main goals is to classify or even reconstruct auditory stimuli based on brain signals of sound perception. Machine learning algorithms work relatively well for speech segments, but perceived—let alone imagined—complex musical stimuli pose a severe challenge. The main focus of this chapter is on deep neural networks—a machine learning technique that uses supervised (i.e., examples of input–output pairs are provided) or unsupervised learning to represent information. The authors discuss two classes of deep neural networks in more detail: convolutional neural networks (CNNs) and recurrent neural networks (RNNs). CNNs show similarities with the actual brain structure of the visual cortex and are good at spatial processing; on the other hand, RNNs are often used to process sequential data. Both classes have been used to estimate musical tempo or classify imagined musical excerpts. In the last section, Ofner and Stober discuss deep generative models, which have the advantage of using unsupervised learning. Although training and analysis of such models are complicated—small deviations (e.g., a shift in tempo) can lead to large mathematical errors—these models are regularly used in real-time applications or interactive settings. One future challenge identified by the authors is to collect more multimodal data, for instance by combining neuroimaging data with physiological, (facial) motor, or behavioural data. This request is in line with Belfi's (this volume) suggestion to make reverse inference—that is, inferring cognitive processes from brain signals—less speculative. Another request is to pair deep neural networks with predictive coding (Clark, 2013) to construct models that are even more similar to the biological brain and that are able to explain internally generated, sensory-like experiences such as mental imagery. All the approaches and methods discussed in this and previous chapters have not only the potential to transform basic research questions related to music and mental imagery but also wide-reaching ethical implications. Some of these will be addressed in the last chapter of Part II.

Informed by his own research in Japan, Papua New Guinea, and Pakistan, Athanasopoulos reflects, in Chapter 10, on methodological and ethical issues in cross-cultural studies of music-related mental imagery. When conducting research in a foreign culture, every concept needs to be revisited before formulating meaningful research questions. Since language is often a barrier, researchers should present images or videos and use non-verbal responses (e.g., free drawings) in their studies. Due to complex (digital) interconnections and dependencies between cultures and societies, researchers should also carefully check their hypotheses and assumptions when doing field work. Ethical considerations should always proceed any research questions, and great attention should be paid to power relations as well as concepts and ideas researchers bring to the field. Athanasopoulos also weaves in findings from his own cross-cultural research, speculating that there may be innate associations between sound and imagery concepts. For data to be interpreted in light of cultural particularities, it is important to form interdisciplinary research teams, including anthropologists, ethnomusicologists, psychologists, and cognitive scientists. Some concrete factors to be taken into account when testing individuals from a different culture encompass formal school experience, musical training, literacy, physical aspects of the experimental setup, familiarity with the type of data collection, and scale construction. Crucially, Athanasopoulos reminds music and mental imagery researchers that "there is no objective quality to the perception of musical imagery; instead, each participant group is likely utilising approaches which maximally facilitate communication between themselves and their peers."

Part III: Mental Imagery and Related States of Consciousness

The third part of this volume explores imagery-based states of consciousness (ordinary or altered) that are evoked during music listening or making. Imaginative thought is remarkably diverse, and beyond the modalities in which it can occur (described in detail in Part I), there are several mental states or states of consciousness that strongly rely on mental imagery (e.g., autobiographical memory, mind-wandering, daydreaming, and absorption) and that can also be relevant to musical contexts. Experts from diverse fields, including psychology, musicology, and music therapy, discuss how music can shape such mental states and underlying processes. The reader may witness consistent overlaps among the different states of consciousness examined here, and the chapters highlight the lack of definite conceptual boundaries among them. In this regard, an outstanding issue for future research on mental imagery and music consists of stimulating conceptual work to provide a holistic framework that can effectively account for the variety of music-related, imagery-based states of consciousness.

In Chapter 11, Jakubowski deals with music-evoked autobiographical memories (MEAMs), which are often listed among the most intense and moving experiences one can have with music. After providing a brief picture of the crucial role of mental imagery in autobiographical memory, Jakubowski specifically focuses on how music can act as an effective cue for spontaneously eliciting positive and vivid lifetime memories. When looking at the peculiarities of MEAMs compared with autobiographical memories evoked by other cues, such as photographs of famous faces and TV programmes, it is striking that the imaginal content of MEAMs exhibits increased motor, spatial, and social elements along with a greater proportion of perceptual details. These findings stress the social nature of music and the important role it plays in creating and nurturing social bonds and the embodied nature of musical experiences. In the last part of her chapter, Jakubowski identifies tasks for future research, including an in-depth investigation of the imagery modalities through which MEAMs occur (given, e.g., the conflicting findings regarding the role of visual imagery), and an exploration of the potential for using music to elicit memories in people with dementia and other memory impairments.

Konishi provides, in Chapter 12, a state-of-the-art review of research on mind-wandering—a ubiquitous mental phenomenon that has long been studied in experimental psychology under a range of different terminologies, such as "daydreaming," "task-unrelated thought," or "stimulus-independent thought." Mind-wandering episodes, which are characterized by a shift of attention away from the task at hand or the external environment, can be investigated in terms of their phenomenological content (very often people mind-wander to current concerns or goals) and can be modulated by the characteristics of the context (mind-wandering increases with fatigue). After reviewing mind-wandering's benefits (e.g., increased creativity) and costs (e.g., decreased task performance), Konishi examines its relationship with music from a two-sided perspective, by addressing the following questions: How can music influence mind-wandering, and can mind-wandering occur in the form of musical imagery? While music seems capable of modulating the content (and to some extent also the frequency) of mind-wandering episodes via emotion, mind-wandering can also occur in the auditory modality, for example, as in the case of *earworms* when a catchy tune plays again and again in our minds. The integration of music into mind-wandering research may provide novel insights into our understanding of conscious mental experiences in relationship to tasks that do not require much focused attention (such as music listening).

In Chapter 13, Gritten offers a theoretical reading of a mental phenomenon closely related to mind-wandering: distraction. Distraction is one mechanism underlying the evocation of mind-wandering that disrupts the focus of the listener's attention, possibly leading to "failed listening." By retrieving a personal anecdote of attending the UK premiere of Goehr's orchestral work *Colossos or Panic*, Gritten unpacks the complex relationship between traction and distraction. Although, at first, distraction may appear as a "drag" on the listener, by triggering mind-wandering and leading away from focused listening, it ultimately challenges the listener. This challenge consists of drawing the listener back to their body and adapting to the materiality of musical sound and to what happens in the external environment. In summary, Gritten's chapter sheds positive light on distraction, showing the functionality it may fulfil and how it consequently amplifies indeterminacy within the listening context.

In Chapter 14, Herbert covers *musical daydreaming*—a heteronomous, multimodal mode of listening marked by fluctuations between internal mentation and awareness of the external sensory environment, which largely overlaps with the concept of mind-wandering. As Herbert explains, heteronomous listening differs from autonomous listening (where music is the sole focus of attention), allowing more space for the unfolding of mental imagery. The chapter is centred around an examination of subjective reports of music listening experiences, which illuminates the phenomenology of musical daydreams as well as how visual imagery contributes to the overall experience by shaping specific content, partly via enculturation through repeated exposure to film. It also clearly emerges that multimodality is a key characteristic of such mental experiences, which are modulated by a rich network of internal/external variables, including age, musical training, context, personality, and intention. Although musical daydreams may strongly rely on biological rhythms, chronobiological explanations remain at present merely speculative, and empirical data are needed to shed more light on this. However, Herbert suggests that this is a highly promising direction for future work, since musical daydreams "may function as a self-regulatory process affording respite from the vicissitudes of daily life—a space for simply 'being.'"

Küssner and Orlandatou (Chapter 15) explore synaesthesia—a condition observed in a small subset of the general population, in which a stimulus not only is processed by its corresponding sense but also activates another sensory modality. The chapter examines synaesthesia's relationship to mental imagery, looking in particular at the comparison of sound-colour synaesthesia and music-induced visual mental imagery. After a brief historical overview of synaesthesia research, the authors put forward the argument that music-induced visual mental imagery can be regarded as a weak form of sound-colour synaesthesia, given the large overlap between their conceptual and experiential features. However, some significant differences also emerge. For example, synaesthesia tends to feature stronger vividness and less control on the experience than music-induced visual mental imagery. The authors underscore three key areas—emotion, creativity, and memory—where synaesthetes' and non-synaesthetes' responses to music-induced visual mental imagery may differ. The chapter ends with the encouragement for future researchers investigating music-induced visual imagery to assess whether their participants are synaesthetes to avoid potential confounds.

With Vroegh's contribution (Chapter 16), we slightly move away from ordinary consciousness and focus on *absorption*, which has often been defined as a trance-like state of consciousness. Aiming to make progress on theoretical debates by adopting an empirical-oriented approach, Vroegh uses a probabilistic graphical approach to unveil how multiple dimensions of consciousness relate to each other in the context of an absorbed listening state, with a particular focus on the dimension of visual imagery. Specifically, he applies Bayesian network analysis on a novel dataset gathered by 193 participants, who, after listening to a 5-minute piece of favourite music, had to fill in the Phenomenology of Consciousness Inventory (Pekala, 1991), which taps onto a number of consciousness dimensions such as attention, self-awareness, short-term memory, self-control, and altered experience. The results show that an absorbed state of mind is linked to a high probability (74%) of experiencing *visual imagery*. Moreover, *altered experience* represents a central "hub" of the network, whereas *mixed affect*, *short-term memory*, and *reflective thoughts* are outcome attributes. Overall, this chapter contributes to uncover the main dimensions characterizing the experience of absorption in music, underscoring its close relationship with visual imagery. It provides intriguing insights into the structure of absorption from a multidimensional perspective on consciousness and puts forward a promising analytic method that could be applied to the study of other music-related and imagery-based states of consciousness.

By combining perspectives from music therapy, shamanic healing, and psychedelic therapy, Fachner explores, in Chapter 17, music-evoked imagery in altered states of consciousness (ASC). The author puts forward the thesis that a recumbent body posture during an ecstatic ASC induces a process of downregulation of arousal that in turn allows to free up energy for focusing on the imagery evoked during music listening. To support his argument, Fachner showcases a range of ASC induction settings including, for example, shamanic journeys and monotonous drumming. The common denominator is that inwardly tuned attention and less movement are used to save up energy, which in turn will be released to boost inner imagery related to the various ASC. Music featuring monotonous structures seems ideal to retreat from the "here and now"; however, the choice of music is culturally mediated, and complex structures may also be relevant in triggering more articulated mental images and narratives. Importantly, Fachner points out that, besides the induction settings, performance rites, suggestibility traits, and personal willingness to enter an altered state play a crucial role in enhancing overall imagery.

Part IV: Applied Mental Imagery

How can mental imagery, which is intangible, only existing in the mind of the person who imagines, have practical applications? Part IV of the current volume addresses this question and, in six chapters, the authors present their research and perspectives on the application of mental imagery. Scholars and practitioners explore a range of domains in applied mental imagery, from motor rehabilitation and coping with music performance anxiety to music therapy and piano performance. Drawing on data from multiple research methods such as interviews, subjective observations, and empirical testing, the authors interpret their findings and the literature while developing reviews and theoretical frameworks. The chapters in Part IV touch on issues which relate back to the previous parts of the volume, underlie our understanding of imagery, and inform the volitional and spontaneous uses of imagery not only in musical settings but also in non-musical settings, thereby responding to calls for translational research in the field.

The first three chapters discuss applications of mental imagery in rehabilitation, therapeutic, and pedagogical settings. In Chapter 18, Schaefer delivers an overview of how imagery is and could be harnessed in motor rehabilitation for movement-related purposes (e.g., in Parkinson's disease) and in music pedagogy for memorization and expressiveness. First, the author explores past research and provides an account of imagery and its intricate link to movement. She touches on issues regarding imagery, its phenomenological aspects, cognitive, behavioural, and neural correlates, and individual imagery abilities. Moreover, she discusses how understanding imagery's underlying processes can improve applications in both rehabilitation and pedagogy. She offers an explanation of how musical imagery (voluntary and spontaneous), similar to music processing, can act as an endogenous auditory cue for movement and could potentially supplement or even replace music in motor rehabilitation. Next, Schaefer reviews research on the use of multimodal imagery in music pedagogy and how various imagery modalities interact with motor-related processes to support memorization of music, reduce cognitive load while performing, enhance communication and expressiveness, and benefit music performance. The commonalities between imagery processes and motor aspects for movement recovery and music pedagogy open up exciting prospects not only for understanding imagery in more depth but also for creating applicable protocols and teaching methods.

Next, moving away from the physical and cognitive benefits of imagery, Finch and Oakman (Chapter 19) present a comprehensive review of the literature on how voluntary imagery is used by musicians to manage affective-related aspects of performance and, more specifically, music performance anxiety (MPA). The authors review intervention studies and provide a categorization of the existing imagery techniques that have been tested with musicians to help them cope with MPA. They employ the superordinate term *mental preparation imagery* comprising four techniques—metaphorical imagery, relaxation imagery, systematic desensitization, and imagery in mental skills training—which are used for general performance reasons, when preparing for specific performances, or for specific performance-related goals. In addition, the authors use the term *mental rehearsal* for certain imagery-related performance tasks. The techniques and the details offered draw on methods from music therapy to cognitive behavioural therapy aimed at reducing anxiety, easing the cognitive and motor demands of the performance, or both. In view of future developments, the authors highlight the strengths and limitations (e.g., small sample size) of the reviewed studies while also making suggestions, based on research results and practices from other disciplines such as sports psychology, about what musical imagery aspects and individual differences to be studied and how.

Imagery techniques for MPA, covered in the previous chapter, adopt aspects from the music therapy form of Guided Imagery and Music (GIM), which is presented in Chapter 20 by Dukić. Although GIM therapy sessions are explained and used as the departure point in this first empirical chapter of Part IV, the reported study utilizes the GIM context as a stimulus to explore one of the hypothesized meanings of music, that is, to offer a narrative. According to Dukić, a *narrative* is a sequence of meaningful events that present a serial order of setup, confrontation, and resolution. This order is realized through active and passive exchanges as expressed in tension and relaxation conveyed by specific features of music acting as a scaffold for imagery to emerge in the same manner. In Dukić's study, participants underwent a standard GIM session and listened to music that was intended to induce (or not) a narrative plot while reporting on spontaneous visual imagery experiences the moment they occurred. The analysis focused on categorization of imagery as passive or active. The results showed that specific musical pieces in the GIM programme, such as Britten's *Sentimental Sarabande*, followed the narrative order as predicted, while others did not. The author offers explanations for these differences, for example, specific musical features that could trigger these responses, and makes suggestions for future studies, for instance, looking at the specific music features, which contribute to the emergence of active and passive imagery. Research on GIM seems promising not only for further understanding imagery and the meaning of music but also, more importantly, for developing tailored music therapy sessions according to the needs of the clients.

The last three chapters cast light on the experience of imagery in skilled music performance. Two chapters deal with accounts of imagery in typical and visually impaired professional pianists and their performances in Western classical music contexts in the UK and South Africa, adapting existing or developing novel theoretical frameworks. In Chapter 21, Presicce presents a brief review of music performers' experiences with visual imagery and provides a system of categories building on a framework of music listening and visual imagery by Taruffi and Küssner (2019). The framework is based on cognitive control constraints on the occurrence of visual and other forms of imagery but this time adapted for performing music. Presicce divides visual imagery during music performance into three categories—*spontaneous*, *heuristic*, and *strategic*—which serve different aspects of performing. She describes how these forms of imagery can enhance the performance, its expressiveness, or its preparation, echoing some of Finch and Oakman's techniques on coping with MPA. In the second part of the chapter, Presicce presents an insider's view and outlines her first-hand experiences as a pianist and member of a piano duo as examples not only to confirm some of the imagery categories of the framework but also to showcase the power of imagery as a performative tool for memorization and expressivity.

In a similar manner, Herbst and van Zyl explore, in Chapter 22, the experience of piano performance and learning but this time by offering a unique window into the inner world of visually impaired (VI) pianists. The sample of four VI pianists (congenitally blind, partially sighted) and one sighted piano teacher with experience in teaching VI pianists provides an insight into a population that has been minimally studied in relation to mental imagery and music. The authors conducted open-ended semi-structured interviews focusing on experiences of listening, performing, and learning music in relation to, for example, the appearance of non-musical associations with music theoretical concepts and mental imagery linked to music notation systems. The identified themes reflect aspects related to VI pianists and how their external experience (e.g., braille notation) could hamper access to music and how their internal experience (e.g., use of metaphors and mind-wandering occurrence) can facilitate their musical understanding and compensate for their visual impairment. Finally, the authors develop a conceptual framework by compiling their findings and interpreting them through the lens of embodied cognition and dynamic system theories. This multimodal framework showcases the synergy of various imagery modalities, affective aspects, music structure, and bodily processes and the practical implications which are relevant for VI and sighted musicians alike.

In Chapter 23, Black presents an intriguingly different view of imagery from what we have seen so far. The chapter provides an examination of how, in the context of an observational study using video recording and interviews, 21 directors of amateur choirs employ verbal content, coined as *verbalized imagery*, to alter the vocal responses of the singers. The definition of verbalized imagery used by the author is distinct from definitions of auditory verbal imagery or musical imagery that have been presented so far in this volume. *Verbalized imagery*, as adopted by Black, refers to the words used orally by the choir directors, which describe images, metaphors, similes, or other figurative language, to elicit or alter singers' vocal responses. Through a very rich dataset and a qualitative analysis, the author identifies several categories of verbalized imagery such as visual, auditory, kinaesthetic, emotional, conceptual-musical, conceptual non-musical, and multimodal. Focusing on visual and emotional verbalized imagery, Black observes tone quality and expression to be the most frequent vocal effects. Given the plethora of types of verbalized imagery and vocal responses, the author encourages choir directors to use verbalized imagery as a strategy in their practice. The use of the term imagery in this chapter is an apt reminder to always ask people what they mean when they talk about "imagery." A clarification of the term will help to avoid misunderstandings not only between scholarly fields concerned with imagery research but also between academia, the arts, and the general public.

Part V: Outlook

As mentioned in the beginning of this introductory chapter, the different definitions of mental imagery provided in this volume are intended to sensitize the reader to the diversity of approaches one may take when studying music and mental imagery, with a view to stimulating more collaborative research on this topic. The issue of defining mental imagery is also taken up by Eerola in the last chapter of this volume. Besides offering his own diagnostic take on the problems to be solved in the near future to make progress in this field such as theoretical integration, analytical reviews, and innovative studies, he re-visits the problem of assessing and measuring mental imagery in musical contexts and proffers an embodied account of mental imagery to integrate existing findings and to be able to deal with complex features such as multimodality.

We envisage that this volume will bring loose ends of mental imagery research together and reveal interconnections between topics and areas of music research that will produce synergistic effects for novel studies. Progress will depend on taking aboard all the theoretical, conceptual, historical, methodological, epistemological, and ethical issues related to music and mental imagery discussed in this volume and elsewhere. It is not an easy task, but one that—if achieved in an interdisciplinary research environment—will potentially transform our knowledge of musical experience and provide a new window onto mental imagery more generally.

References

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204. https://doi.org/10.1017/S0140525X12000477

Connolly, C., & Williamon, A. (2004). Mental skills training. In A. Williamon (Ed.), *Musical excellence: Strategies and techniques to enhance performance* (pp. 221–245). Oxford: Oxford University Press.

Day, R. A., & Thompson, W. F. (2019). Measuring the onset of experiences of emotion and imagery in response to music. *Psychomusicology: Music, Mind, and Brain*, *29*(2–3), 75–89. https://doi.org/10.1037/pmu0000220

Floridou, G. A., Williamson, V. J., Stewart, L., & Müllensiefen, D. (2015). The Involuntary Musical Imagery Scale (IMIS). *Psychomusicology: Music, Mind, and Brain*, *25*(1), 28–36. https://doi.org/10.1037/pmu0000067

Gelding, R. W., Thompson, W. F., & Johnson, B. W. (2019). Musical imagery depends upon coordination of auditory and sensorimotor brain activity. *Scientific Reports*, *9*(1), 16823. https://doi.org/10.1038/s41598-019-53260-9

Godøy, R. I. (2019). Thinking sound-motion objects. In M. Filimowicz (Ed.), *Foundations in sound design for interactive media* (pp. 161–178). New York: Routledge.

Godøy, R. I., & Jørgensen, H. (Eds.). (2001). *Musical imagery*. New York: Taylor & Francis.

Gracyk, T. (2019). Imaginative listening to music. In M. Grimshaw-Aagaard & M. Walther-Hansen (Eds.), *The Oxford handbook of sound and imagination* (Vol. 2). Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190460242.013.11

Halpern, A. A., & Overy, K. (2019). Voluntary auditory imagery and music pedagogy. In M. Grimshaw-Aagaard, M. Walther-Hansen, & M. Knakkergaard (Eds.), *The Oxford handbook of sound and imagination* (Vol. 2, pp. 390–407). Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190460242.013.49

Hashim, S., Stewart, L., & Küssner, M. B. (2020). Saccadic eye-movements suppress visual mental imagery and partly reduce emotional response during music listening. *Music & Science*, *3*, 2059204320959580. https://doi.org/10.1177/2059204320959580

James, W. (1890). *The principles of psychology*. New York: Henry Holt and Co.

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, *31*(5), 559–575. https://doi.org/10.1017/S0140525X08005293

Killingly, C., Lacherez, P., & Meuter, R. (2021). Singing in the brain: Investigating the cognitive basis of earworms. *Music Perception*, *38*(5), 456–472. https://doi.org/10.1525/mp.2021.38.5.456

Kosslyn, S. M., Behrmann, M., & Jeannerod, M. (1995). The cognitive neuroscience of mental imagery. *Neuropsychologia*, *33*(11), 1335–1344. https://doi.org/10.1016/0028-3932(95)00067-D

Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006). *The case for mental imagery*. Oxford: Oxford University Press.

Liikkanen, L. A., & Jakubowski, K. (2020). Involuntary musical imagery as a component of ordinary music cognition: A review of empirical evidence. *Psychonomic Bulletin & Review, 27*(6), 1195–1217. https://doi.org/10.3758/s13423-020-01750-7

Liptak, T. A., Omigie, D., & Floridou, G. A. (2022). The idiosyncrasy of Involuntary Musical Imagery Repetition (IMIR) experiences: The role of tempo and lyrics. *Music Perception: An Interdisciplinary Journal*, *39*(3), 320–338. https://doi.org/10.1525/mp.2022.39.3.320

Nanay, B. (2018). Multimodal mental imagery. *Cortex*, *105*, 125–134. https://doi.org/10.1016/j.cortex.2017.07.006

Nanay, B. (2021). Unconscious mental imagery. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1817), 20190689. https://doi.org/10.1098/rstb.2019.0689

Pekala, R. J. (1991). *Quantifying consciousness: An empirical approach*. New York: Plenum Press.

Schaefer, R. S. (2017). Music in the brain: Imagery and memory. In R. Ashley & R. Timmers (Eds.), *The Routledge companion to music cognition* (pp. 25–35). New York: Routledge.

Schaefer, R. S., Vlek, R. J., & Desain, P. (2011). Music perception and imagery in EEG: Alpha band effects of task and stimulus. *International Journal of Psychophysiology*, *82*(3), 254–259. https://doi.org/10.1016/j.ijpsycho.2011.09.007

Schneider, A., & Godøy, R. I. (2001). Perspectives and challenges of musical imagery. In R. I. Godøy & H. Jørgensen (Eds.), *Musical imagery* (pp. 5–26). New York: Taylor & Francis.

Taruffi, L., & Küssner, M. B. (2019). A review of music-evoked visual mental imagery: Conceptual issues, relation to emotion, and functional outcome. *Psychomusicology: Music, Mind, and Brain*, *29*(2–3), 62–74. https://doi.org/10.1037/pmu0000226

Thomas, N. J. T. (1999). Are theories of imagery theories of imagination? An active perception approach to conscious mental content. *Cognitive Science*, *23*(2), 207–245. https://doi.org/10.1207/s15516709cog2302_3

Waller, D., Schweitzer, J. R., Brunton, J. R., & Knudson, R. M. (2012). A century of imagery research: Reflections on Cheves Perky's contribution to our understanding of mental imagery. *The American Journal of Psychology*, *125*(3), 291–305. https://doi.org/10.5406/amerjpsyc.125.3.0291

Watanabe, H., Tanaka, H., Sakti, S., & Nakamura, S. (2020). Synchronization between overt speech envelope and EEG oscillations during imagined speech. *Neuroscience Research*, *153*, 48–55. https://doi.org/10.1016/j.neures.2019.04.004

Wilain, S., Omigie, D., Küssner, M. B., Taruffi, L., & Floridou, G. A. (2021, July 28–31). *The Multimodal Mental Imagery of Music Scale (MMIMS): A novel instrument for measuring individual differences in multimodal mental imagery during music listening* [Paper presentation]. 16th International Conference on Music Perception and Cognition and the 11th Triennial Conference of the European Society for the Cognitive Sciences of Music, Sheffield, UK.

Part I

Modalities of Mental Imagery

1 The Chronicles of Musical Imagery as It Occurs Before, During, and After Music

Georgia A. Floridou

A large part of our everyday sensory experience is auditory. Close your eyes, and you will become aware of the external sonic environment and its richness in voices, nature sounds, city sounds, and music. Close your ears, and you'll find the sonic richness of the external world continuing internally in your mind in the form of auditory mental imagery, that is, "sounds in the mind." One of the most powerful external and internal auditory experiences is that of music. Music experience can be generated externally, during perception of music when sound waves reach our eardrums, or internally, during musical mental imagery. In comparison to other imagery modalities (e.g., visual and motor), the external music experience has been mostly studied in relation to its internal counterpart, musical imagery. However, imagery in other modalities also contributes to our musical experience; for example, speech imagery (Taruffi et al., 2017) and visual imagery (Taruffi & Küssner, 2019, this volume) can also occur while listening to music. In this chapter, I focus on the most studied modality of imagery in relation to music, which is at the very core of the musical experience, musical imagery (MI henceforth).

The most common operational definition of MI is the following: MI is "the experience of imagining musical sound in the absence of directly corresponding sound stimulation from the physical environment" (Bailes, 2007, p. 555). Prima facie this definition serves its purpose, however, I highlight two issues related to it which stem from the most up-to-date research advancements in this area. First, it assumes that the experience is unimodal and solely musical. When imaging music, non-auditory modalities can also be present. Kalakoski (2001) argued that MI stands at the intersection of several sensory (stimulus) modalities, such as visual, motor, and kinaesthetic (i.e., the feeling of muscle movement). For example, motor imagery can also be present during musical imagery not only due to the intrinsic relation between music and physical action (e.g., music performance and dance) but also because of the role of time as a key link between sound and motor processing (Bishop, 2018; Schaefer, 2014, Godøy, this volume). Secondly, while "the absence of directly corresponding sound stimulation" (Bailes, 2007, p. 555) indeed encapsulates the majority of the cases where direct external stimulation is absent and MI occurs offline (i.e., when not perceiving but still processing the corresponding musical stimulus), the aforementioned statement ignores the idea that,

DOI: 10.4324/9780429330070-3

as I will explain later, MI can also occur online, that is, in the presence of directly corresponding music stimulation.

Musical Imagery Before, During, and After Music

Internally and externally generated music experiences unfold over time. Here, I synthesize diverse empirical findings about the internally generated music experience in a temporal fashion when it occurs in relation to the externally generated music experience. I conceptualize and introduce a novel framework using a tripartite classification about the occurrence of the various MI forms at three temporal points in relation to music: before, during, and after music (e.g., composing, playing, and listening). Within the framework and for each MI form at each temporal point, I present the *what* (i.e., the defining features of the MI forms), the *who* (i.e., the relation of MI with musicianship), and the *why* (i.e., the functions MI serves), as well as the link between MI forms and other cognitive systems, for example, creative and spontaneous cognition, predictive processing, and embodied cognition. Reviewing MI through the lens of such a framework allows us to see the heterogeneity of MI based on variation of its elements, which contribute to the diverse functional outcomes that MI offers, despite the apparently similar basis. The framework presented (see Table 1.1) facilitates a better understanding[1] of a seemingly simple everyday manifestation of thought, which yet reflects the complex and multifaceted workings of the mind.

Musical Imagery Before Music

Two instances of MI occurring before the composition as well as the playing of new music have been described in the literature—novel MI and voluntary MI, respectively. The first form, novel MI, is associated with creative thinking and consequently could aid musical composition. Composers experience the genesis of a tune in their minds, voluntarily (i.e., wilfully and deliberately) or spontaneously (i.e., without intention and involuntarily), repeatedly or not, during wakefulness or sleep. Research on the role of novel MI in musical creativity is extremely limited and exclusively based on self-reports (Bailes, 2006; Floridou, 2016). What we do know so far is that, during wakefulness, musicians, but also non-musicians, experience novel MI (Bailes, 2006, 2007; Beaman & Williams, 2010; Beaty et al., 2013), something that demonstrates that musical creative thinking is not exclusive to musicians. In an exploratory study, I conducted interviews with composers who experience novel involuntary and repeated MI that could lead to compositional output (Floridou, 2016). The study revealed the context in which such an MI form can come to mind, namely low attention states and repetitive movements, states which also apply to familiar MI, spontaneous as well as creative forms of cognition (Baird et al., 2012; Williamson et al., 2012). In addition, the musical features of novel MI reflect similarities not only to familiar involuntary but also to voluntary MI, where the melody, rhythm, and timbre are the most prevalently experienced musical features (Bailes, 2007; Halpern, 2007). The

Table 1.1 Conceptual framework about the occurrence of the various forms of musical imagery (MI) at three temporal points in relation to music (composition, playing, and listening) before, during, and after. Each MI form is presented in each temporal point in relation to music along with what features define it, who experiences it, and why it happens.

	Before Music		*During Music*		*After Music*		
	Composition	*Playing*	*Listening*	*Playing*	*Playing/Listening*		
MI Form	*Novel MI*	*Voluntary MI*	*Constructive Imagery*	*Anticipatory Imagery*	*Mind-Pops*	*Earworms*	*Everyday Voluntary MI*
What?	—Spontaneous/ Voluntary —Repetitive/ Non-repetitive —Wakefulness/ Sleep (musical dreams)	—Voluntary —Wakefulness	—Automatic —Wakefulness	—Voluntary —Wakefulness	—Spontaneous —Non-repetitive —Wakefulness	—Spontaneous —Repetitive —Wakefulness	—Wakefulness —Voluntary
Who?	—Musicians —Composers —Non-musicians	—Musicians	—Musicians —Non-musicians	—Musicians	—Musicians —Non-musicians	—Musicians —Non-musicians	—Musicians —Non-musicians
Why?	—Creative thinking	—Mental rehearsal —Learning —Memorization	—Expectations —Affective reactions	—Performance support —Action planning —Accurate timing			

study findings also suggested that the repetitive character of MI could be associated with the functionality of MI, as it was attributed to memory consolidation and its role as a driving force for composition. Finally, the findings suggested that novel MI is based on re-synthesis of previous and familiar experiences and novel combinations, rather than on purely novel musical elements that have never been encountered before.

Musical dreams, which occur during sleep states (König & Schredl, 2019; Uga et al., 2006), are another rare manifestation of novel MI, which have not been studied much. Novel MI in the form of musical dreams is experienced by both musicians and non-musicians, although significantly more by the former, and, compared to other kinds of dreams (non-musical), musical dreams are more pleasant (König & Schredl, 2019). However, beyond anecdotal reports of composers attributing the inspiration for their composition to musical dreams (Grace, 2012) and one study reporting the occurrence of novel MI in the dreams of musicians (Uga et al., 2006), more research is needed to explore the content and the function of novel MI during sleep states.

Once new music is composed, musicians mentally practice in preparation of the music performance. The second MI form occurring before music, voluntary MI, commonly coined as mental rehearsal, is one of the most practical uses of MI for musicians, and is linked with the stages of enhancing the learning of a musical piece. One of the strategies used in this context is notational audiation (Gordon, 1999, 2003), which requires the presence of visual cues in the form of a musical score which is read and represented in the mind in the form of MI. In the absence of external visual cues such as scores or an instrument, MI can still be used for mental rehearsal and can be multimodal (i.e., experienced simultaneously or in succession with other imagery modalities). For example, motor imagery and visual imagery can be used for mentally playing the instrument and setting the scene (e.g., audience and venue). Mentally practicing music using voluntary MI aids musicians while learning for memorization (Davidson-Kelly et al., 2015) and is used as a rehearsal tool before a performance (Clark et al., 2011). In addition, the benefits of such MI include helping to overcome technical difficulties, gaining confidence and fluidity of movement, connecting with an audience, and the avoidance of physical fatigue (Connolly & Williamon, 2004; Davidson-Kelly et al., 2015; Trusheim, 1991). Taken together, these findings suggest that MI serves multiple functional outcomes for musicians and, although not as practical for non-musicians, it is still an indication about their capacity for creative thinking.

Musical Imagery During Music

Can you listen to or perform music and simultaneously imagine it? The answer to this question lies in two online MI forms—constructive imagery and anticipatory imagery, respectively. Constructive imagery, commonly referred to as musical expectations, takes place during music and involves predictive processing, which is a core phenomenon of music cognition. Researchers have put forward the idea that music listening, in musicians and non-musicians alike, often includes an automatic imagery-like component which contributes to making sense of the incoming information by generating expectations in the form of constructive imagery for different aspects of the music, for example, the next note coming (Janata, 2001). Musical expectations have also been attributed a role for explaining affective (i.e., emotional) reactions to music listening (Huron, 2006). In particular, listeners' expectations are not always immediately satisfied, but they might be temporarily delayed. Such violations, disruptions, and resolutions of expectations may subsequently lead to meaningful and expressive moments in music. Furthermore, constructive imagery may also be present during voluntary MI, offering an explanation for the commonalities and shared cerebral processing between perception and imagery of music (Schaefer, 2014).

When performing music, the voluntary use of MI by musicians, along with motor and visual imagery, has been described and studied as anticipatory imagery. Such a form of MI satisfies multiple purposes related to informing the performance and supporting action planning (Keller & Appel, 2010; Keller & Koch, 2006; Keller, 2012). Anticipatory MI also helps musicians with the production of the specific sound they want to reproduce, achieving attention on what is coming next, predictions of their own and their co-performers' action, and accurate timing of movements, ultimately influencing ensemble performance (Pecenka & Keller, 2009). This is an indication that external and internal music experiences are intertwined and that anticipatory MI, along imagery in other modalities, serves multiple purposes in musicians, during music playing. Overall, the findings suggest that constructive MI helps musicians and non-musicians alike to process and make sense of external musical experiences.

Musical Imagery After Music

The final stage of MI occurrence is after music listening and performance.[2] The main offline MI forms that have been investigated are everyday musical mind-pops, involuntary musical imagery repetition, and voluntary MI. Despite these being the most common manifestations of MI in non-musicians but also very frequent experiences in musicians, research on involuntary and voluntary MI in everyday waking life has started only recently but has accelerated similar to analogous research on other forms of everyday cognition such as mind-wandering (Smallwood, 2013), autobiographical memory (Berntsen, 1998), future thinking (Berntsen, 2019), and semantic memory (Kvavilashvili & Mandler, 2004).

Research on musical mind-pops, which are single involuntary occurrences of MI, although scarce (Kvavilashvili & Anthony, 2012; Kvavilashvili & Mandler, 2004), has revealed that they are activated after music exposure (priming) and that only a small percentage of them transform to repeated forms of MI. The most studied form of MI is involuntary musical imagery repetition (IMIR; Liptak et al., 2022) also known as "having an earworm." The core characteristics of IMIR are its involuntary onset and repetitive continuation either in a loop or intermittently

(Floridou et al., 2015). Around 90% of people globally report that they have such an experience at least once a week (Liikkanen, 2012; Liikkanen et al., 2015), each lasting on average 30 minutes (Beaman & Williams, 2010; McNally-Gagnon, 2015). Despite negative stereotypes associated with IMIR, the majority of people experience it as neutral or pleasant (Beaman & Williams, 2010; Floridou & Müllensiefen, 2015; Halpern & Bartlett, 2011).

A plethora of studies has revealed what kind of music induces IMIR, who experiences it, and under what kind of conditions they occur. Research on the music itself has revealed that the experience is highly idiosyncratic, as there is little or no overlap in the tunes experienced as IMIR between or within people (Beaman & Williams, 2010; Hyman et al., 2013; Jakubowski et al., 2015; Liikkanen, 2012; Williamson et al., 2014). Looking specifically at the musical features of reported IMIR, one study has argued that melodic contour patterns and possibly fast tempo are important for their induction (Jakubowski et al., 2017). However, these features interact with extra-musical features such as tune popularity and exposure (Jakubowski et al., 2017). The first empirical study on the topic revealed that the music-listening habits of the individual, in terms of tempo and lyrics, are associated with relevant IMIR occurrence rather than tempo and lyrics per se (Liptak et al., 2022). Yet, more empirical research directly testing musical features is needed, since the findings are based on self-reports.

The next most popular question is who is more susceptible to IMIR? Studies on demographic factors such as gender and age indicate that women (Liikkanen, 2012; McCullough Campbell & Margulis, 2015) and young adults (Floridou et al., 2019; Liikkanen, 2012) experience it more frequently. Personality traits are most commonly associated with the phenomenological characteristics of IMIR and least with their frequency (Floridou et al., 2012), with, for example, people who score higher in Neuroticism perceiving IMIR as more annoying. Engagement with music-related activities seems to be important and responsible for increased IMIR frequency. For example, engagement with music at a professional level (e.g., students and professionals rehearsing or performing) and in everyday life (e.g., listening to music, singing, dancing, or tapping along; Floridou et al., 2012, 2015; McCullough Campbell & Margulis, 2015) have been associated with increased IMIR frequency. The link between IMIR and motor-related behaviours such as the one mentioned earlier has been interpreted and discussed within the framework of embodied cognition (Bailes, 2019), where experiences of the mind are intertwined with experiences of the body.

Finally, what are the conditions which are associated with IMIR occurrence? IMIR is most commonly triggered by exposure to music (recent or repeated), words, sounds, or pictures (Williamson et al., 2012). The context preceding its occurrence is mostly associated with activities such as travelling and exercising, that is, monotonous or repetitive activities which do not require much attention and are characterized by low cognitive load (Floridou et al., 2017; Hyman et al., 2013), similar to other forms of spontaneous cognition (Berntsen et al., 2013; Kvavilashvili & Mandler, 2004).

Everyday MI can also occur voluntarily in musicians and non-musicians, yet relevant research is sparse. Some studies have explored everyday voluntary MI conjointly with involuntary MI forms, without distinguishing between the two (Bailes, 2006, 2007, 2015; Beaty et al., 2013), while others have measured both experiences separately (Cotter et al., 2018; Jakubowski et al., 2018). Comparisons to IMIR have revealed that everyday voluntary MI is less common (Cotter et al., 2018), and it elicits less intense emotional reactions (Jakubowski et al., 2018).

In conclusion, a large body of research suggests that everyday occurrences of involuntary MI are extremely common in musicians and non-musicians alike, but still a number of questions remain and more research is needed. Directly comparing everyday involuntary and voluntary MI forms, as well as musical mind-pops and IMIR, could help to identify differences in relation to the features of intentionality and repetition as well as their functionality which is currently unknown.

Concluding Remarks

MI is a heterogeneous, cross-modal experience similar to music perception. Considering the complexity of MI and all the forms it takes before, during, and after music highlights the extremely idiosyncratic personal profile of MI, which is reflected in our own chronicles of listening, playing, and composing music. Despite MI being the most researched imagery modality in relation to music, it is evident that there are still many questions that require further investigation. Exciting future avenues of research include comparative studies of the different forms MI can take, especially in relation to the cognitive mechanisms involved in their processing. A deeper understanding of the cognitive and neural underpinnings of MI can help harness its power and can contribute towards the development of applications not only for musicians and pedagogical settings for boosting creativity and enhancing learning but also for non-musicians in health and clinical settings (Schaefer, this volume) for motor rehabilitation (Satoh & Kazuhara, 2008; Schaefer et al., 2014) and interventions to treat anxiety (Ulor et al., 2018). MI is at the core of the musical experience and can reveal not only what music is but also what it means to be a musical human being.[3]

Notes

1 To fully understand MI, we also need to consider its *atypology* too, that is, pathological and non-pathological forms of MI where its features are amplified (e.g., musical obsessions and hallucinations) or the inability to form all or some MI features (similar to aphantasia, i.e., individuals who cannot voluntary form visual mental images; Zeman et al., 2015). Atypical cases of MI can give us a glimpse into the phenomenologically different inner world of these populations and an indication about specific neural correlates of MI. However, these cases of MI are not within the scope of this review (for reviews, see Williams, 2015).

2 The boundaries between the temporal points in the framework under discussion are not strict and should not be seen as mutually exclusive but in some cases as a dynamic

interplay. For example, MI occurring before and after music could be the same; however, the distinction lies in the familiarity with the music experienced as MI, the intentionality level, and the voluntariness and goal-directedness before music as compared to the involuntariness and apparent non-goal-directedness after music.

3 G.A.F. was supported by a British Academy Postdoctoral Fellowship (pf160109). I would like to thank Dr Kelly Jakubowski for feedback on an earlier version of the manuscript.

References

Bailes, F. A. (2006). The use of experience-sampling methods to monitor musical imagery in everyday life. *Musicae Scientiae*, *10*(2), 173–190. https://doi.org/10.1177/102986490601000202

Bailes, F. A. (2007). The prevalence and nature of imagined music in the everyday lives of music students. *Psychology of Music*, *35*(4), 555–570. https://doi.org/10.1177/0305735607077834

Bailes, F. A. (2015). Music in mind? An experience sampling study of what and when, towards an understanding of why. *Psychomusicology: Music, Mind, and Brain*, *25*, 58–68. http://dx.doi.org/10.1037/pmu0000078

Bailes, F. A. (2019). Musical imagery and the temporality of consciousness. In R. Herbert, D. Clarke, & E. Clarke (Eds.), *Music and consciousness 2: Worlds, practices, modalities* (pp. 271–285). Oxford: Oxford University Press.

Baird, B., Smallwood, J., Mrazek, M. D., Kam, J. W., Franklin, M. S., & Schooler, J. W. (2012). Inspired by distraction: Mind-wandering facilitates creative incubation. *Psychological Science*, *23*(10), 1117–1122. doi:10.1177/0956797612446024

Beaman, C. P., & Williams, T. I. (2010). Earworms (stuck song syndrome): Towards a natural history of intrusive thoughts. *British Journal of Psychology*, *101*(4), 637–653. doi:10.1348/000712609x479636

Beaty, R. E., Burgin, C. J., Nusbaum, E. C., Kwapil, T. R., Hodges, D. A., & Silvia, P. J. (2013). Music to the inner ears: Exploring individual differences in musical imagery. *Consciousness and Cognition*, *22*(4), 1163–1173. doi:10.1016/j.concog.2013.07.006

Berntsen, D. (1998). Voluntary and involuntary access to autobiographical memory. *Memory*, *6*(2), 113–141. doi: 10.1080/741942071

Berntsen, D. (2019). Spontaneous future cognitions: An integrative review. *Psychological Research*, *83*(4), 651–665. https://doi.org/10.1007/s00426-018-1127-z

Berntsen, D., Staugaard, S. R., & Sørensen, L. M. T. (2013). Why am I remembering this now? Predicting the occurrence of involuntary (spontaneous) episodic memories. *Journal of Experimental Psychology: General*, *142*(2), 426. doi: 10.1037/a0029128

Bishop, L. (2018). Collaborative musical creativity: How ensembles coordinate spontaneity. *Frontiers in psychology*, *9*, 1285.

Clark, T., Williamon, A., & Aksentijevic, A. (2011). Musical imagery and imagination: The function, measurement, and application of imagery skills for performance. In D. Hargreaves, D. Miell, & R. MacDonald (Eds.), *Musical imaginations: Multidisciplinary perspectives on creativity, performance and perception* (pp. 351–368). Oxford: Oxford University Press.

Connolly, C., & Williamon, A. (2004). Mental skills training. In A. Williamon (Ed.), *Musical excellence: Strategies and techniques to enhance performance* (pp. 221–245). Oxford: Oxford University Press.

Cotter, K. N., Christensen, A. P., & Silvia, P. J. (2018). Understanding inner music: A dimensional approach to musical imagery. *Psychology of Aesthetics, Creativity, and the Arts*, *13*(4), 489–503. http://dx.doi.org/10.1037/aca0000195

Davidson-Kelly, K., Schaefer, R. S., Moran, N., & Overy, K. (2015). "Total inner memory": Deliberate uses of multimodal musical imagery during performance preparation. *Psychomusicology: Music, Mind, and Brain*, *25*(1), 83–92. https://doi.org/10.1037/pmu0000091

Floridou, G. A. (2016). *Investigating the relationship between involuntary musical imagery and other forms of spontaneous cognition* [Doctoral Dissertation, Goldsmiths, University of London]. Goldsmiths University Research Online. https://research.gold.ac.uk/19424/1/PSY_thesis_FloridouG_2016.pdf

Floridou, G. A., Halpern, A., & Williamson, V. (2019). Age-related changes in everyday forms of involuntary and voluntary cognition. *PsyArXiv*. doi:10.31234/osf.io/6e4ch

Floridou, G. A., & Müllensiefen, D. (2015). Environmental and mental conditions predicting the experience of involuntary musical imagery: An experience sampling method study. *Consciousness and Cognition*, *33*, 472–486. doi:10.1016/j.concog.2015.02.012

Floridou, G. A., Williamson, V. J., & Müllensiefen, D. (2012). Contracting earworms: The roles of personality and musicality. In E. Cambouropoulos, C. Tsougras, P. Mavromatis, & K. Pastiadis (Eds.), *Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music* (pp. 516–518). Thessaloniki: Aristotle University of Thessaloniki.

Floridou, G. A., Williamson, V. J., & Stewart, L. (2017). A novel indirect method for capturing involuntary musical imagery under varying cognitive load. *The Quarterly Journal of Experimental Psychology*, *70*(11), 2189–2199. https://doi.org/10.1080/17470218.2016.1227860

Floridou, G. A., Williamson, V. J., Stewart, L., & Müllensiefen, D. (2015). The Involuntary Musical Imagery Scale (IMIS). *Psychomusicology: Music, Mind, and Brain*, *25*(1), 28–36. https://doi.org/10.1037/pmu0000067

Gordon, E. E. (1999). All about audiation and music aptitudes: Edwin E. Gordon discusses using audiation and music aptitudes as teaching tools to allow students to reach their full music potential. *Music Educators Journal*, *86*(2), 41–44. doi:10.2307/3399589

Gordon, E. E. (2003). *Learning sequences in music: Skill, content, and patterns: A music learning theory*. Chicago: GIA Publications.

Grace, N. (2012). Music and dreams. In D. Barrett & P. McNamara (Eds.), *Encyclopedia of sleep and dreams: The evolution, function, nature, and mysteries of slumber* (pp. 430–432). Santa Barbara, CA: Greenwood.

Halpern, A. (2007). Commentary on "Timbre as an elusive component of imagery for music," by Freya Bailes. *Empirical Musicology Review*, *2*(1), 35–37.

Halpern, A. R., & Bartlett, J. C. (2011). The persistence of musical memories: A descriptive study of earworms. *Music Perception*, *28*(4), 425–432. doi:10.1525/mp.2011.28.4.425

Huron, D. B. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge, MA: MIT Press.

Hyman, I. E., Jr., Burland, N. K., Duskin, H. M., Cook, M. C., Roy, C. M., McGrath, J. C., & Roundhill, R. F. (2013). Going Gaga: Investigating, creating, and manipulating the song stuck in my head. *Applied Cognitive Psychology*, *27*(2), 204–215. doi:10.1002/acp.2897

Jakubowski, K., Bashir, Z., Farrugia, N., & Stewart, L. (2018). Involuntary and voluntary recall of musical memories: A comparison of temporal accuracy and emotional responses. *Memory & Cognition*, *46*(5), 741–756. https://doi.org/10.3758/s13421-018-0792-x

Jakubowski, K., Farrugia, N., Halpern, A. R., Sankarpandi, S. K., & Stewart, L. (2015). The speed of our mental soundtracks: Tracking the tempo of involuntary musical imagery in everyday life. *Memory & Cognition*, *43*(8), 1229–1242. https://doi.org/10.3758/s13421-015-0531-5

Jakubowski, K., Finkel, S., Stewart, L., & Müllensiefen, D. (2017). Dissecting an earworm: Melodic features and song popularity predict involuntary musical imagery. *Psychology of Aesthetics, Creativity, and the Arts*, *11*(2), 122–135. https://doi.org/10.1037/aca0000090

Janata, P. (2001). Brain electrical activity evoked by mental formation of auditory expectations and images. *Brain Topography*, *13*(3), 169–193. https://doi.org/10.1023/A:1007803102254

Kalakoski, V. (2001). Musical imagery and working memory. In R. Godøy & H. Jørgensen (Eds.), *Musical imagery* (pp. 43–55). New York, NY: Taylor & Francis.

Keller, P. E. (2012). Mental imagery in music performance: Underlying mechanisms and potential benefits. *Annals of the New York Academy of Sciences*, *1252*(1), 206–213. doi:10.1111/j.1749-6632.2011.06439.x

Keller, P. E., & Appel, M. (2010). Individual differences, auditory imagery, and the coordination of body movements and sounds in musical ensembles. *Music Perception*, *28*(1), 27–46. doi:10.1525/mp.2010.28.1.27

Keller, P. E., & Koch, I. (2006). The planning and execution of short auditory sequences. *Psychonomic Bulletin & Review*, *13*(4), 711–716. doi:10.3758/BF03193985

König, N., & Schredl, M. (2019). Music in dreams: A diary study. *Psychology of Music*, *49*(3), 351–359. https://doi.org/10.1177/0305735619854533

Kvavilashvili, L., & Anthony, S. (2012, November 15–18). *When do Christmas songs pop into your mind? Testing a long-term priming hypothesis* [Poster presentation]. Annual Meeting of Psychonomic Society, Minneapolis, Minnesota.

Kvavilashvili, L., & Mandler, G. (2004). Out of one's mind: A study of involuntary semantic memories. *Cognitive Psychology*, *48*(1), 47–94. doi:10.1016/S0010-0285(03)00115-4

Liikkanen, L. A. (2012). Musical activities predispose to involuntary musical imagery. *Psychology of Music*, *40*(2), 236–256. doi:10.1177/0305735611406578

Liikkanen, L. A., Jakubowski, K., & Toivanen, J. M. (2015). Catching earworms on Twitter: Using big data to study involuntary musical imagery. *Music Perception: An Interdisciplinary Journal*, *33*(2), 199–216. doi:10.1525/mp.2015.33.2.199

Liptak, T., Omigie, D., & Floridou, G. A. (2022). The idiosyncrasy of Involuntary Musical Imagery Repetition (IMIR) experiences: The role of tempo and lyrics. *Music Perception: An Interdisciplinary Journal, 39*(3), 320–338.

McCullough Campbell, S., & Margulis, E. H. (2015). Catching an earworm through movement. *Journal of New Music Research*, *44*(4), 347–358. https://doi.org/10.1080/09298215.2015.1084331

McNally-Gagnon, A. (2015). *Imagerie musicale involontaire: Caractéristiques phénoménologiques et mnésiques* [Doctoral Dissertation, University of Montreal]. https://papyrus.bib.umontreal.ca/xmlui/handle/1866/16051

Pecenka, N., & Keller, P. (2009). Auditory pitch imagery and its relationship to musical synchronization. *Annals of the New York Academy of Sciences*, *1169*(1), 282–286. doi:10.1111/j.1749-6632.2009.04785.x

Satoh, M., & Kuzuhara, S. (2008). Training in mental singing while walking improves gait disturbance in Parkinson's disease patients. *European Neurology*, *60*(5), 237–243. https://doi.org/10.1159/000151699

Schaefer, R. S. (2014). Images of time: Temporal aspects of auditory and movement imagination. *Frontiers in Psychology*, *5*, 877. https://doi.org/10.3389/fpsyg.2014.00877

Schaefer, R. S., Morcom, A. M., Roberts, N., & Overy, K. (2014). Moving to music: Effects of heard and imagined musical cues on movement-related brain activity. *Frontiers in Human Neuroscience*, *8*, 774. https://doi.org/10.3389/fnhum.2014.00774

Smallwood, J. (2013). Distinguishing how from why the mind wanders: A process—occurrence framework for self-generated mental activity. *Psychological Bulletin*, *139*(3), 519. https://doi.org/10.1037/a0030010

Taruffi, L., & Küssner, M. B. (2019). A review of music-evoked visual mental imagery: Conceptual issues, relation to emotion, and functional outcome. *Psychomusicology: Music, Mind, and Brain*, *29*(2–3), 62–74. https://doi.org/10.1037/pmu0000226

Taruffi, L., Pehrs, C., Skouras, S., & Koelsch, S. (2017). Effects of sad and happy music on mind-wandering and the default mode network. *Scientific Reports*, *7*(1), 14396. doi:10.1038/s41598-017-14849-0

Trusheim, W. H. (1991). Audiation and mental imagery: Implications for artistic performance. *Quarterly Journal of Music Teaching and Learning*, *2*(1–2), 139–147.

Uga, V., Lemut, M. C., Zampi, C., Zilli, I., & Salzarulo, P. (2006). Music in dreams. *Consciousness and Cognition*, *15*(2), 351–357. doi:10.1016/j.concog.2005.09.003

Ulor, M., Bailes, F., & O'Connor, D. B. (2018, July, 23–28). *Using voluntary musical imagery as an intervention for anxiety: How can this be studied?* [Poster presentation]. 15th International Conference on Music Perception and Cognition/10th Triennial Conference of the European Society for the Cognitive Sciences of Music, Graz, Austria. https://static.uni-graz.at/fileadmin/veranstaltungen/music-psychology-conference2018/documents/ICMPC15ESCOM10abstractbook.pdf

Williams, T. I. (2015). The classification of involuntary musical imagery: The case for earworms. *Psychomusicology: Music, Mind and Brain*, *25*(1), 5–13. https://doi.org/10.1037/pmu0000082

Williamson, V. J., Jilka, S. R., Fry, J., Finkel, S., Mullensiefen, D., & Stewart, L. (2012). How do "earworms" start? Classifying the everyday circumstances of involuntary musical imagery. *Psychology of Music*, *40*(3), 259–284. https://doi.org/10.1177/0305735611418553

Williamson, V. J., Liikkanen, L. A., Jakubowski, K., & Stewart, L. (2014). Sticky tunes: How do people react to involuntary musical imagery? *PLoS One*, *9*(1), e86170. https://doi.org/10.1371/journal.pone.0086170

Zeman, A., Dewar, M., & Della Sala, S. (2015). Lives without imagery—Congenital aphantasia. *Cortex*, *73*, 378–380. http://dx.doi.org/10.1016/j.cortex.2015.05.019

2 Visual Mental Imagery, Music, and Emotion

From Academic Discourse to Clinical Applications

Liila Taruffi and Mats B. Küssner

In this chapter, we examine music-related visual mental imagery, which is evident across many cultures, in different forms and functions. We highlight some of its features and particularities, and discuss two academic debates in more depth: the format of such mental representations and their role in music-induced emotions. In the last section, we argue that particularly the link to emotions makes music-related visual mental imagery an interesting candidate for clinical interventions.

According to Kosslyn et al. (1995, p. 1335), "[v]isual mental imagery is 'seeing' in the absence of the appropriate immediate sensory input." *Music-related visual mental imagery*—as a broad umbrella term—can thus be described in the most general sense as "seeing" in the absence of the appropriate immediate sensory input while listening to, or making, music. Many of us have experienced visual mental imagery (VMI) during music listening—whether such mental representations consist of memories from the past, learned associations, simple abstract shapes, or complex visual narratives (e.g., listening to instrumental music from the romantic era resulting in imagining a natural landscape). Music seems capable to foster the onset of VMI, or to provide the context for its unfolding, and to modulate its features,[1] making it somewhat controllable. Music-related VMI may co-occur and share properties with mental imagery evoked through other modalities. For example, imagining a choreography while listening to music may require the co-occurrence of visual and motor imagery (see Godøy, this volume), both of which must be synchronized with the musical material. In this sense, VMI may unfold as part of a holistic, multimodal phenomenon (see the concept of "musical daydreams" in Herbert, this volume), significantly increasing the level of complexity and variables through which we can examine VMI.[2]

Forms, Functions, and Features of Music-Related Visual Mental Imagery Around the Globe

The specific facets, dimensions, and mechanisms of music-related VMI are determined by both psychological and socio-cultural constraints (for a more detailed conceptualization, see also Taruffi & Küssner, 2019). Although it can be assumed that VMI has been experienced in all cultures throughout human history, forms and functions of VMI during music listening (and making) may differ significantly.

DOI: 10.4324/9780429330070-4

For instance, Siberian shamans receive special training to enhance their VMI skills, as the ability to cultivate visions is a crucial component of their culture's music-based healing rituals (Siikala, 1978). The aesthetics of Indian classical music are deeply rooted in VMI (Leante, 2009), as illustrated by cultural items such as *rāgmālā* (i.e., miniature paintings that depict Indian musical modes called ragas; Ebeling, 1973). In Western societies, passive music therapy approaches such as Guided Imagery and Music (GIM; Bonny, 2002; Dukić, this volume) employ standardized playlists of classical music to stimulate VMI and trigger spontaneous associations; VMI can help improve the therapist's understanding of the client's subconscious emotional processes and may shed additional light on the client's personal issues. Importantly, VMI may be one reason for music's emotional power (Juslin & Västfjäll, 2008). VMI has often emotional connotations that are likely to result in the creation of intricate networks of cross-modal visual, emotional, and auditory components. The precise nature of these cross-modal relationships, specifically (a) their causal direction and (b) the mapping of each modality on the others (e.g., dark colours/negative emotions/dark timbres), is yet to be determined. In non-therapeutic musical settings, choral directors and musicians commonly harness VMI during rehearsals in order to accomplish desired interpretations or sound qualities (see chapters by Black, this volume; Presicce, this volume), while film music composers often create music that is strongly suggestive of particular images or atmospheres in order to enhance the film narrative (MacDonald, 2013).

As the aforementioned examples show, VMI does not appear in a standardized manner across cultures and individuals: it may vary in frequency, quality, duration, and vividness. VMI may also vary greatly across the temporal domain, and may belong to the past, present, future, or even be timeless (such as in cases of abstract imagery; e.g., moving shapes). Depending on the context, VMI may be deliberate or spontaneous, and, in the latter case, overlap with "mind-wandering" (see Konishi, this volume). VMI exhibits during various levels of consciousness, spanning from normal waking states to altered states of consciousness, where the components of arousal, awareness, and vigilance are modulated (see Fachner, this volume). For instance, absorption—an aspect of *trance* characterized by an effortless involvement with the object of experience (see Vroegh, this volume)—is typically accompanied by strong and vivid VMI both during episodes of everyday "spontaneous trancing" with music (Herbert, 2011) and during music rituals in ethnic ceremonies (Rouget, 1985). Thus, music's capacity to shape VMI makes it an invaluable tool for application to various therapeutic, educational, and performance contexts.

Visual Mental Imagery in Cognitive Sciences and Philosophy: The Format of Mental Representations

The colloquial expression "seeing with the mind's eye" suggests that VMI closely resembles the perceptual experience of vision. Along these lines, cognitive psychologists define VMI as mental representations that preserve the perceptive properties of visual stimuli and ultimately elicit the subjective experience of vision. From a mechanistic standpoint, this definition suggests that, beyond its content, the *form* in which VMI is encoded is depictive (i.e., similar to a drawing or photograph). Still, the format (or type of code) of these mental representations is likely to be very complex. In fact, the issue of VMI's format has been one of the most debated topics in cognitive psychology and philosophy since the 1970s and has been strongly influenced by evidence from neuroscience in more recent years, though its theoretical basis still remains controversial. One of the main reasons behind the so-called "imagery debate" (Sterelny, 1986) is that VMI is a fundamentally *private* phenomenon and is therefore difficult to objectively assess. The "internal" nature of VMI led many behavioural psychologists in the mid-20th century to deny its existence and condemn any scientific investigation of it (Thomas talks about "iconophobia" among the behaviourists; Thomas, 2008, section 1.1). In this regard, John B. Watson took the most radical position, stigmatizing imagery as an esoteric notion and maintaining that it is only a sort of verbally mediated memory (Watson, 1928).[3]

It has been proposed that mental images can be *propositional* or *depictive* (Kosslyn et al., 2006). Propositional representations are abstract language-like representations close to the language in formal logic. For example, a propositional representation of a girl eating ice cream could be: EAT (GIRL, ICE-CREAM). Conversely, depictive representations would be exemplified as a drawing of a girl eating ice cream. The first format type makes explicit and accessible semantic interpretations (i.e., a different symbol is used for ambiguous words), while the latter makes *explicit* and *accessible* all aspects of shape and relations between shape and other perceptual qualities such as colour (Marr, 1982). As described in detail by Kosslyn et al. (2006), a fundamental turn in the aforementioned debate occurred with the advent of neuroimaging data, which allowed for a more direct investigation of VMI's neural basis and also for this basis to be compared with that of visual perception. Several areas of the human visual cortex are organized topographically, meaning that adjacent locations in the retina correspond to adjacent locations in the visual cortex. This also entails that the distance between two points on the cortex *depicts* the distance between the corresponding points in space (Kosslyn et al., 2006). Moreover, local damage to topographic areas of the cortex produces localized scotoma (i.e., a blind spot) in the corresponding part of the retina, indicating that the topographic properties of these areas play a role in vision. Crucially, fMRI studies were able to show that such topographic areas of the visual cortex are engaged—in a similar yet not identical manner—during both perception and imagery (see, e.g., the meta-analysis by Kosslyn et al., 2006). This implication that the topographic organization of the brain may be used both for perception and for imagination of an external reality speaks strongly in favour of the depictive claims underlying VMI.[4] With regard to music research, a study (Koelsch et al., 2013) demonstrating activation of visual areas during listening to fear-evoking music not only suggests that music-induced VMI taps similar neural processes as general VMI but also highlights its relevance for affective responses. Similarly, subsequent studies provided more evidence of the engagement of the primary visual cortex during listening to emotional music with eyes closed (e.g., Taruffi et al., 2021).

Visual Mental Imagery and Music: The Relationship With Affect

Because our experiences with music often involve affective responses, emotions may in turn strongly impact the quality and the outcome of evoked VMI. But the opposite may also be true: VMI could elicit and/or shape emotional responses. So far, most empirical research conducted on the topic of VMI and music has focused on its relationship to emotion. In this section, we will provide a summary of this strand of research, highlighting general trends and outstanding issues.

The focus on the link between emotion and VMI can be explained by the fact that VMI has been theoretically proposed as a mechanism underlying music-evoked emotion (Juslin & Västfjäll, 2008). A number of studies suggest that VMI and emotional processes during music listening are closely intertwined. For instance, while investigating whether sad and happy music are linked to different levels of self-reported mind-wandering and differential recruitment of brain regions, Taruffi et al. (2017) provided evidence that mind-wandering evoked during music listening is predominantly a visual phenomenon. In particular, their results showed that sad music was associated with enhanced mind-wandering and stronger activity within the main nodes of the brain's Default Mode Network (DMN), which is the principal neural contributor to mind-wandering (Mason et al., 2007), and has also been proposed to be involved in aesthetic experiences of visual art (Belfi et al., 2019; Vessel et al., 2019, but note that there is no evidence available that links the DMN uniquely to the visual modality). Regarding the form of mental experiences (i.e., whether spontaneous thoughts occur in the form of visual imagery or inner language), VMI was by far the dominant modality for both sad and happy music conditions, suggesting that mind-wandering during music listening overlaps with VMI regardless of the emotion experienced during the music. In line with findings from Taruffi et al. (2017), mind-wandering literature has provided robust evidence for the role of affect in shaping the quantity and quality of mind-wandering episodes (for a review, see Konishi, this volume). While higher rates of mind-wandering tend to lead to negative mood (Killingsworth & Gilbert, 2010; Smallwood et al., 2007), it is important to notice that this association is strongly mediated by the content of thoughts and/or images and that not all mind-wandering is detrimental for mood. Specifically, past-related thoughts/images are strongly associated with higher levels of unhappiness (Ruby et al., 2013; Smallwood & O'Connor, 2011). Martarelli et al. (2016) manipulated the valence of music (i.e., by using sad and happy instrumental music excerpts) to investigate the mediating role of daydreams on relaxation and music liking. Although they did not directly assess the presence of VMI in participants' daydreams,[5] they found that daydreams that arose during the musical experience mediated the effect of the music's type (i.e., sad vs. happy) on relaxation and liking. Specifically, sad music correlated with increased relaxation, whereas happy music promoted more positive daydreams, which in turn facilitated relaxation and were associated with participants' increased liking of music. These findings imply that *daydreams*—mental phenomena that seem to be dominated by the visual modality—play a functional role during music listening. In summary, the emotional tone of music appears to be linked to the content of associated thought (e.g., happy music is associated with happier thoughts compared with sad music) as corroborated further by a recent study using heroic and sad music (Koelsch et al., 2019). Nevertheless, it remains to be tested directly whether the relationship between thoughts/images and music's affective tone is mediated by the valence and/or arousal of the music rather than by its overall emotional quality (e.g., does sad or slow music lead to increased mind-wandering?). Further experiments should disentangle the effect of arousal and/or valence from single emotions on VMI, perhaps by including music conditions that are similar in dimensional categories (e.g., anger and happy music, which have high arousal but differing valence). A recent study (Herff et al., 2021) has provided further evidence of systematic effects of music, including tempo, on image features. One hundred participants engaged in a directed imagination task that required closing their eyes and imagining a continuation of a journey towards a mountain that was displayed in a short video clip, while listening to different music pieces or silence. After the imagined journeys, participants reported vividness, the imagined time passed and distance travelled, and the emotional tone of the imagined content. The results showed that all the main variables were influenced by the music pieces and, in particular, faster tempo predicted less imagined time passed and distance travelled, as well as more positive sentiment (Herff et al., 2021).

Further support for the importance of VMI during affective processes in musical contexts has been provided by experience sampling and survey studies (Juslin et al., 2008; Küssner & Eerola, 2019). In Juslin et al.'s study (2008), college students carried a palmtop device that emitted a sound signal several times per day at random intervals for two weeks. When signalled, participants were required to complete a self-report questionnaire that asked whether they had experienced a music-evoked emotion and, if so, what its cause had been. This approach allowed the researchers to explore emotions in music as they naturally occurred in everyday life. Results showed that VMI was rated as the fourth most commonly reported cause of emotion (7%), following emotional contagion (32%), brain stem responses such as a loud or sudden sound evoking tension or surprise (25%), and episodic memory (14%). Findings from this study support the involvement of VMI in emotion evocation, although this does not seem to be the primary mechanism through which emotions are evoked by music. While exploring the prevalence, nature, and functions of VMI in music and how it is associated with domain-specific skills by assessing individual differences in musical sophistication, Küssner and Eerola (2019) found that one of the main functions of VMI during music listening is modulation of emotional arousal. Factor analyses of a novel music-related visual imagery questionnaire resulted in two components, labelled *Vivid Visual Imagery* and *Soothing Visual Imagery*, being identified among both musically trained and untrained listeners, suggesting the use of VMI for calming or relaxing functions, whereas only musically trained individuals reported using imagery to feel more energized and excited. Furthermore, musically untrained individuals reported experience of vivid visual imagery, though not for use in explicitly modulating emotional responses.

One outstanding issue emerging from this literature is that the assessment of mental images and emotions took place simultaneously, making it impossible to understand whether VMI precedes or follows emotion. Day and Thompson (2019) examined the temporal relationship between VMI and emotions evoked during music listening using response times. In their experiment, participants were asked to indicate when they (a) recognized an emotion in music, (b) experienced an affective response to music, and (c) experienced a visual image related to music. Results (which however apply to ~90% of participants) showed that the time taken to *perceive* an emotion was significantly shorter than the time taken to *feel* an emotion, which was, in turn, significantly shorter than the time taken to *experience* VMI. These findings are particularly interesting, as they stand in contrast to Juslin and Västfjäll's (2008) framework on VMI as a mechanism of emotion evocation. However, the causal relationship between VMI and emotions may be modulated by the quality of the emotional responses as suggested by Vroegh (2018), where, in an online survey, individuals were asked to listen to music excerpts varying in valence and arousal and to provide information regarding their attentional focus, frequency, and vividness of VMI, and quality of their felt emotions. The best-fitting model accounting for the data was characterized by a unidirectional connection from emotion to VMI when explicitly positive emotions were concerned; however, in the case of mixed emotions, the results pointed to a reverse direction, from VMI to felt emotions.

While the studies reviewed in this section support the importance and the complexity of the relationship between felt emotions and VMI in musical contexts, novel empirical paradigms encompassing more objective and sophisticated assessments should be employed in order to provide a more reliable picture of this phenomenon (for an extensive discussion on methodological issues in VMI research, please refer to Part II of this volume). Self-reports—employed by the majority of the studies mentioned earlier—may not be the optimal index of VMI due to the phenomenon's ephemeral nature likely enhancing memory bias (for a review, see Pylyshyn, 2002). Measuring VMI is a challenging task and requires researchers to consider a wider range of additional measures to triangulate self-reports, including physiological (e.g., eye-tracking), neuroimaging (e.g., electroencephalography and functional magnetic resonance imaging), and indirect markers (e.g., response times). Some progress on measuring as well as suppressing VMI has been made in clinical contexts.

Visual Mental Imagery in Clinical Psychology: Potential Music-Based Interventions

A comprehensive understanding of how music shapes VMI is likely to spur new opportunities for clinical treatment and intervention. Dysfunctional VMI occurs in a wide range of mental disorders. For example, intrusive imagery is a distinctive feature of post-traumatic stress disorder; hallucinatory imagery is observed in schizophrenia; and intrusive memories of negative past events are typically found in depression (Pearson et al., 2013). Furthermore, the relationship between VMI and emotions is crucial in clinical settings, since VMI often works as an emotional amplifier (Holmes et al., 2008), as has been evinced for the specific case of anxiety (Holmes & Mathews, 2005). Interestingly, this "increasing" effect of VMI on emotion has also been corroborated for certain aesthetic emotions, such as sublimity and unease, in the context of a live opera performance (Balteș & Miu, 2014). In summary, VMI represents a potential threat to clinical populations that typically experience negative affect. This begs the question: how could music be harnessed in therapy? As mentioned previously, GIM is a viable existing technique within music therapy that harnesses spontaneous generation of VMI in order to reduce stress (Beck et al., 2015) and chemotherapy-induced anxiety or vomiting (Karagozoglu et al., 2013). Furthermore, music could also be incorporated into existing imagery-focused techniques in order to modulate unwanted, intrusive, and maladaptive VMI (see also Herff et al., 2021). Examples of such techniques are cognitive behavioural therapy (involving repeated imaginary exposure to a feared object in patients with anxiety disorders; Foa et al., 1980), imagery re-scripting (featuring the switch of negative to positive imagery; Holmes et al., 2007), systematic desensitization (aiming to remove the fear response of a phobia, and to substitute a relaxation response to the conditional stimulus gradually using counter conditioning; Wolpe, 1958), and eye movement desensitization and reprocessing (triggering lateral eye movements during the recall of emotional memories to diminish their vividness; Chen et al., 2014; van den Hout et al., 2012). The interaction between auditory and visual modalities should be carefully considered to test the effectiveness of music. In this regard, Hashim et al. (2020) provide evidence that an eye movement distractor task can suppress VMI during music listening and lead to attenuated emotional responses. Still, further research is needed to clarify the interaction between music and VMI and to consider potential differential effects of various parameters such as the music's acoustic or stylistic features (e.g., genre) and interindividual differences among listeners (e.g., familiarity or preference).

Conclusion

VMI is an important component of the human experience of music listening and making. The emotional power of VMI in musical contexts has been recognized and harnessed by numerous cultures around the world, manifesting itself in different practices and interventions ranging from shamanic rituals to chemotherapy. We have summarized current scientific findings and prominent theoretical accounts regarding the underlying mechanisms of music-related VMI, particularly with respect to emotional response. However, to unlock the full potential of music-related VMI in clinical and other applied contexts, a deeper understanding of the causal relationship among VMI, music, and emotion is yet to be achieved.

Notes

1 This is likely to occur partly through associations with the musical and acoustic material.
2 Certainly, multimodal imagery applies to non-musical contexts as well. However, musical experiences are intrinsically multifaceted and very often involve all our sensory modalities (both perception and imagery). For this reason, music represents an extremely effective tool to investigate the phenomenon of multimodal imagery.
3 Although it is difficult to identify a robust reason underlying behaviourists' strong denial of VMI—besides their exclusive interest in empirically observable behaviours—anecdotal evidence suggested that Watson might have had strong *auditory* but weak *visual* imagery and Faw (2009) even pointed to the possibility that poor VMI abilities may have characterized a few behaviourist psychologists.
4 However, this does not exclude the possibility of a hybrid system (i.e., neurons are not only arranged in a functional shape but also code for specific properties).
5 Evidence for this is nevertheless provided by the study of Taruffi et al. (2017), which also made use of instrumental sad and happy music.

References

Balteș, F. R., & Miu, A. C. (2014). Emotions during live music performance: Links with individual differences in empathy, visual imagery, and mood. *Psychomusicology: Music, Mind, and Brain*, *24*(1), 58. https://doi.org/10.1037/pmu0000030

Beck, B. D., Hansen, Å. M., & Gold, C. (2015). Coping with work-related stress through Guided Imagery and Music (GIM): Randomized controlled trial. *Journal of Music Therapy*, *52*(3), 323–352. https://doi.org/10.1093/jmt/thv011

Belfi, A. M., Vessel, E. A., Brielmann, A., Isik, A. I., Chatterjee, A., Leder, H., Pelli, D. P., & Starr, G. G. (2019). Dynamics of aesthetic experience are reflected in the default-mode network. *NeuroImage*, *188*, 584–597. https://doi.org/10.1016/j.neuroimage.2018.12.017

Bonny, H. L. (2002). *Music & consciousness: The evolution of guided imagery and music*. Gilsum: Barcelona Publishers.

Chen, Y. R., Hung, K. W., Tsai, J. C., Chu, H., Chung, M. H., Chen, S. R., . . . Chou, K. R. (2014). Efficacy of eye-movement desensitization and reprocessing for patients with posttraumatic-stress disorder: A meta-analysis of randomized controlled trials. *PLoS One*, *9*(8), e103676. https://doi.org/10.1371/journal.pone.0103676v

Day, R. A., & Thompson, W. F. (2019). Measuring the onset of experiences of emotion and imagery in response to music. *Psychomusicology: Music, Mind, and Brain*, *29*(2–3), 75–89. https://doi.org/10.1037/pmu0000220

Ebeling, K. (1973). *Ragamala painting*. Basel: Ravi Kumar.

Faw, B. (2009). Conflicting intuitions may be based on differing abilities: Evidence from mental imaging research. *Journal of Consciousness Studies*, *16*(4), 45–68.

Foa, E. B., Steketee, G., Turner, R. M., & Fischer, S. C. (1980). Effects of imaginal exposure to feared disasters in obsessive-compulsive checkers. *Behaviour Research and Therapy*, *18*(5), 449–455. https://doi.org/10.1016/0005-7967(80)90010-8

Hashim, S., Stewart, L., & Küssner, M. B. (2020). Saccadic eye-movements suppress visual mental imagery and partly reduce emotional response during music listening. *Music & Science*, *3*, 1–10. https://doi.org/10.1177%2F2059204320959580

Herbert, R. (2011). Consciousness and everyday music listening: Trancing, dissociation, and absorption. In D. Clarke & E. Clarke (Eds.), *Music and consciousness: Philosophical, psychological, and cultural perspectives* (pp. 295–308). Oxford: Oxford University Press.

Herff, S. A., Cecchetti, G., Taruffi, L., & Déguernel, K. (2021). Music influences vividness and content of imagined journeys in a directed visual imagery task. *Scientific Reports*, *11*(1), 1–12. https://doi.org/10.1038/s41598-021-95260-8

Holmes, E. A., Arntz, A., & Smucker, M. R. (2007). Imagery rescripting in cognitive behaviour therapy: Images, treatment techniques and outcomes. *Journal of Behavior Therapy and Experimental Psychiatry*, *38*(4), 297–305. https://doi.org/10.1016/j.jbtep.2007.10.007

Holmes, E. A., Geddes, J. R., Colom, F., & Goodwin, G. M. (2008). Mental imagery as an emotional amplifier: Application to bipolar disorder. *Behaviour Research and Therapy*, *46*(12), 1251–1258. https://doi.org/10.1016/j.brat.2008.09.005

Holmes, E. A., & Mathews, A. (2005). Mental imagery and emotion: A special relationship? *Emotion*, *5*(4), 489–497. https://doi.org/10.1037/1528-3542.5.4.489

Juslin, P. N., Liljeström, S., Västfjäll, D., Barradas, G., & Silva, A. (2008). An experience sampling study of emotional reactions to music: Listener, music, and situation. *Emotion*, *8*(5), 668–683. https://doi.org/10.1037/a0013505

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, *31*(5), 559–575. doi:10.1017/S0140525X08005293

Karagozoglu, S., Tekyasar, F., & Yilmaz, F. A. (2013). Effects of music therapy and guided visual imagery on chemotherapy-induced anxiety and nausea—vomiting. *Journal of Clinical Nursing*, *22*(1–2), 39–50. https://doi.org/10.1111/jocn.12030

Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science*, *330*(6006), 932–932. https://doi.org/10.1126/science.1192439

Koelsch, S., Bashevkin, T., Kristensen, J., Tvedt, J., & Jentschke, S. (2019). Heroic music stimulates empowering thoughts during mind-wandering. *Scientific Reports*, *9*(1), 1–10. https://doi.org/10.1038/s41598-019-46266-w

Koelsch, S., Skouras, S., Fritz, T., Herrera, P., Bonhage, C., Küssner, M. B., & Jacobs, A. M. (2013). The roles of superficial amygdala and auditory cortex in music-evoked fear and joy. *Neuroimage*, *81*, 49–60. https://doi.org/10.1016/j.neuroimage.2013.05.008

Kosslyn, S. M., Behrmann, M., & Jeannerod, M. (1995). The cognitive neuroscience of mental imagery. *Neuropsychologia*, *33*(11), 1335–1344. https://doi.org/10.1016/0028-3932(95)00067-D

Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006). *The case for mental imagery*. New York: Oxford University Press.

Küssner, M. B., & Eerola, T. (2019). The content and functions of vivid and soothing visual imagery during music listening: Findings from a survey study. *Psychomusicology: Music, Mind, and Brain*, *29*(2–3), 90–99. https://doi.org/10.1037/pmu0000238

Leante, L. (2009). The lotus and the king: Imagery, gesture and meaning in a Hindustani rāg. *Ethnomusicology Forum*, *18*(2), 185–206. https://doi.org/10.1080/17411910903141874

MacDonald, L. E. (2013). *The invisible art of film music: A comprehensive history*. Lanham: Scarecrow Press.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York: W. H. Freeman.

Martarelli, C. S., Mayer, B., & Mast, F. W. (2016). Daydreams and trait affect: The role of the listener's state of mind in the emotional response to music. *Consciousness and Cognition*, *46*, 27–35. https://doi.org/10.1016/j.concog.2016.09.014

Mason, M. F., Norton, M. I., Van Horn, J. D., Wegner, D. M., Grafton, S. T., & Macrae, C. N. (2007). Wandering minds: The default network and stimulus-independent thought. *Science*, *315*(5810), 393–395. https://doi.org/10.1126/science.1131295

Pearson, D. G., Deeprose, C., Wallace-Hadrill, S. M., Heyes, S. B., & Holmes, E. A. (2013). Assessing mental imagery in clinical psychology: A review of imagery measures and a guiding framework. *Clinical Psychology Review*, *33*(1), 1–23. https://doi.org/10.1016/j.cpr.2012.09.001

Pylyshyn, Z. W. (2002). Mental imagery: In search of a theory. *Behavioral and Brain Sciences*, *25*(2), 157–182. doi:10.1017/S0140525X02000043

Rouget, G. (1985). *Music and trance: A theory of the relations between music and possession*. Chicago: University of Chicago Press.

Ruby, F. J., Smallwood, J., Engen, H., & Singer, T. (2013). How self-generated thought shapes mood—the relation between mind-wandering and mood depends on the socio-temporal content of thoughts. *PLoS One*, *8*(10), e77554. https://doi.org/10.1371/journal.pone.0077554

Siikala, A. L. (1978). The rite technique of the Siberian shaman. *FF Communications Turku*, *93*(220), 3–385.

Smallwood, J., & O'Connor, R. C. (2011). Imprisoned by the past: Unhappy moods lead to a retrospective bias to mind wandering. *Cognition & Emotion*, *25*(8), 1481–1490. https://doi.org/10.1080/02699931.2010.545263

Smallwood, J., O'Connor, R. C., Sudbery, M. V., & Obonsawin, M. (2007). Mind-wandering and dysphoria. *Cognition & Emotion*, *21*(4), 816–842. https://doi.org/10.1080/02699930600911531

Sterelny, K. (1986). The imagery debate. *Philosophy of Science*, *53*(4), 560–583. https://doi.org/10.1086/289340

Taruffi, L., & Küssner, M. B. (2019). A review of music-evoked visual mental imagery: Conceptual issues, relation to emotion, and functional outcome. *Psychomusicology: Music, Mind, and Brain*, *29*(2–3), 62–74. https://doi.org/10.1037/pmu0000226

Taruffi, L., Pehrs, C., Skouras, S., & Koelsch, S. (2017). Effects of sad and happy music on mind-wandering and the default mode network. *Scientific Reports*, *7*(1), 1–10. https://doi.org/10.1038/s41598-017-14849-0

Taruffi, L., Skouras, S., Pehrs, C., & Koelsch, S. (2021). Trait empathy shapes neural responses toward sad music. *Cognitive, Affective, & Behavioral Neuroscience*, *21*(1), 231–241. https://doi.org/10.3758/s13415-020-00861-x

Thomas, N. (2008). Imagery. In E. N. Zalta (Ed.), *Stanford encyclopedia of philosophy*. Stanford University. http://plato.stanford.edu/entries/mental-imagery/.

van den Hout, M. A., Rijkeboer, M. M., Engelhard, I. M., Klugkist, I., Hornsveld, H., Toffolo, M. J., & Cath, D. C. (2012). Tones inferior to eye movements in the EMDR treatment of PTSD. *Behaviour Research and Therapy*, *50*(5), 275–279.

Vessel, E. A., Isik, A. I., Belfi, A. M., Stahl, J. L., & Starr, G. G. (2019). The default-mode network represents aesthetic appeal that generalizes across visual domains. *PNAS*, *116*(38), 19155–19164. https://doi.org/10.1073/pnas.1902650116

Vroegh, T. (2018, May 16–19). *Investigating the directional link between music-induced visual imagery and two qualitatively different types of emotional responses* [Poster presentation]. KOSMOS Workshop "Mind Wandering and Visual Mental Imagery in Music", Berlin, Germany.

Watson, J. B. (1928). *The ways of behaviourism*. New York: Harper & Brothers.

Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition*. Stanford: Stanford University Press.

3 Intermittent Motor Control in Volitional Musical Imagery

Rolf Inge Godøy

How can composers, arrangers, or conductors predict sound features of music just by reading scores? Or how can musicians rehearse music in their minds without recourse to an instrument? These questions concern what we call *triggers of volitional musical imagery*, and although in the last couple of decades we have seen a large number of publications on musical imagery topics (including on so-called *involuntary musical imagery*, e.g., Williams, 2015), we seem not to have so many publications on what are efficient and robust triggers of volitional musical imagery (see also Floridou, this volume).

Yet during these same decades, research on music-related body motion (e.g., Godøy & Leman, 2010) has documented that mental images of sound are closely related to mental images of sound-producing body motion, suggesting that mental images of body motion could actually trigger mental images of sound. The basis for such trigger relationships is the massive amounts of previous experiences most people have of causality in sound production and which I have tried to summarize in the concept of *motormimetic cognition* (Godøy, 2001, 2003): I can imagine hitting a drum, and thus conjure up the "boom" sound of the drum, or I can imagine my fist softly depressing bass region keys on the piano, and with that, the sound of a soft and dark piano cluster chord.

Accepting that mental images of musical sound may be triggered by mental images of sound-producing motion, the next question becomes how to trigger mental images of sound-producing body motion. Of current theories on motion initiation in human movement science, the most suitable for our purposes seems to be the theory of so-called *intermittent motor control* (e.g., Karniel, 2013; Loram et al., 2014; Sakaguchi et al., 2015), a theory claiming that human motor control proceeds in a discontinuous, point-by-point, and anticipatory manner, because continuous feedback control would be too slow for skilled fast body motion as is typically required in music performance. Musicians need to be ahead of events, that is, to, in an instant, visualize and prepare for the upcoming sound-producing body motion trajectories. In other words, musicians need to adopt *piecewise anticipatory motor control* and *point-by-point motion initiation*, so in this chapter, I will argue that volitional musical imagery can be triggered by similar intermittent schemes in motor imagery.

DOI: 10.4324/9780429330070-5

The aim of this chapter is then to present what I like to call the "intermittency hypothesis" for musical imagery, and some of the research findings and conceptual arguments that converge in suggesting that this is indeed a plausible hypothesis. That said, the prospect of developing concrete techniques for reliable volitional musical imagery could potentially also be of practical benefit for performers, composers, and improvisers, and for anyone interested in exploring their own mental images of music.

Object-Based Musical Imagery

Musical imagery may encompass several different features, and we have seen studies that focus on imagery for pitch, melodies, timbre, dynamics, timing, and expression (see Cotter, 2019, for an updated overview). But the content of musical imagery could also be explored in view of holistic chunks of musical sound, hence, as more in line with our current topic of intermittency. Such a chunk-focused approach was in fact advocated by Pierre Schaeffer and co-workers as the foundation of a more universal music theory based on *sound objects*, that is, on chunks typically in the 0.5–5 seconds duration range (Schaeffer, 1966, 1967/1998). Interestingly, we may in so-called ecological psychology (Gaver, 1993), as well as in more recent work on imagery for multimodal events (Nanay, 2018), see focusing on sonic events that resemble the holistic sound objects of Schaeffer. However, Schaeffer's sound objects are based on their overall dynamical shapes, rather than their source identity and/or everyday significance.

Schaeffer's ideas of sound objects grew out of the use of sound fragments in the *musique concrète*, and evolved from being basically a collage composition tool into becoming the basis for a new music theory. This theory comprised a scheme for the top-down ordering of salient subjective perceptual features in music, extending from the images of overall dynamic, pitch-related, and timbre-related envelopes in what was called the *typology of sound objects* (Schaeffer, 1966, pp. 429–459), down to images of internal timbral-textural detail features called the *morphology of sound objects* (Schaeffer, 1966, pp. 509–597). One of the remarkable features of this scheme is that of capturing so many of the most salient features as *shapes*, and as will be argued later, shapes that are inherently holistic and lend themselves very well to intermittent motor control.

The typology has three main categories of dynamic shapes—*impulsive*, *sustained*, and *iterative*—and three main categories of pitch-related features—*stable*, *variable*, and *unpitched* (inharmonic or noise-dominated sounds). The morphology is extensive with numerous categories and sub-categories for the internal features of sound objects, and just to mention two prominent ones, *grain* (the rapid fluctuation of dynamic, pitch, or timbre in a sound object, e.g., as in a tremolo) and *gait* (the slower undulating motion in a sound object, typically at a walking or dancing pace). Equally important is that these sound object shapes resemble body motion shapes, shapes that we experience in synchrony with sounds, for example, a protracted violin sound as similar to a protracted hand motion shape, an impulsive drum sound as similar to an abrupt hand motion shape, and a piano tremolo sound as similar to a rapid back-and-forth shaking hand motion shape. This means that we may have motormimetic components in parallel with the sound object features, hence what can be called *sound-motion objects* (Godøy, 2019).

Motor-Related Musical Imagery

Following Schaeffer's approach of "interroger la conscience qui écoute" (Schaeffer, 1966, p. 147, in English: "questioning the listening consciousness") and the top-down strategy of digging further into the subjectively perceived features of sound objects, we will see that the sound object and its features are not about "pure sound," but are composite, in particular, so that the border between sound patterns and body motion patterns may become blurred (Godøy, 2006). Once we make this transition to the multimodal, we can see how sound objects are shaped by sound-producing motion, and how motion features become part of the sound objects, firstly in the observable features, that is, visual images of motion trajectories and postures (using video, motion capture, or just seeing and annotating), and secondly, in more indirect observations of effort, for example, based on motion velocity and electromyographic (EMG) data of muscle activation, see, for example, Gonzalez Sanchez et al. (2019). The motion elements all contribute to sound object formation, evident in the following *audible features*:

- All kinds of contours, compared with the *typology* mentioned earlier, manifest in rhythmical-textural figures, as well as in motives, melodic fragments, ornaments, and so forth are *de facto* also body motion shapes.
- All kinds of timbral and harmonic shapes, compared with the *morphology* mentioned earlier, are also reflected in body motion shapes.

And upon a closer study of sound-producing body motion, we can list some crucial *motor control features* that contribute to sound object formation:

- *Phase-transition*, denoting a shift in grouping based on rate and proximity of motion elements, for example, the transition from a slowly pulsating sound to a tremolo sound because of increase in the rate of sound onsets (Godøy, 2014).
- *Coarticulation*, the fusion of motion elements because of the transitions between postures, resulting in a contextual smearing of singular events (Godøy, 2014). In fact, performance and interpretation are coarticulatory fusions of singular tone events into coherent phrases, that is, into coherent sound-motion objects.
- *Hierarchies of motor control* so that certain tasks are carried out automatically and without detail control (Grafton & Hamilton, 2007; Klapp & Jagacinski, 2011).

- The *psychological refractory period* (PRP) denoting a temporal bottleneck in motor control at around approximately 500 milliseconds (Klapp & Jagacinski, 2011; Loram et al., 2014).
- *Action gestalts*, that is, preprogrammed motor control units without need for continuous control, only needing intermittent control (Klapp & Jagacinski, 2011).
- *Postures* at salient moments in time as frames of reference for body motion (Rosenbaum et al., 2007; Rosenbaum, 2017).
- In addition, we have *motion optimization* elements such as shifts between effort and relaxation, conservation of momentum, and exploitations of rebounds that point in the direction of intermittent effort.

But needless to say, we still have substantial challenges in exploring various features of sound-producing body motion. In particular, issues of motion initiation (including intermittent motor control) seem to be not well researched, and we are currently working along three main avenues here:

- indications of intermittency in the motion data (Sakaguchi et al., 2015)
- indications of intermittency in the EMG signals (Aoki et al., 1989)
- with co-workers, reverse engineering of intermittent control in relation to continuous motion

However, the main point for now is to see how motormimetic simulation of sound-producing motion makes for salient presence of sound-motion objects in musical imagery.

Shape Cognition

What emerges from reflections on musical sound and music-related body motion is that most features of both sound and body motion may be envisaged as instances of shapes:

- postures and motion trajectory shapes of the sound-producing effectors (fingers, hands, arms, vocal apparatus, etc.)
- shapes of the derivatives of sound-producing motion, that is, of velocity, acceleration, and jerk of the effectors and of other motion data, for example, EMG, as shapes

And we have the following shapes in perceived sound, often similar to the shapes of sound-producing motion:

- contour shapes of dynamics, pitch, and timbre of single sounds
- shapes of motives, textures, rhythmical patterns, tempi, and expressivity in more composite sounds and sound passages

When it comes to shapes in musical imagery, it seems to be not so well researched, but besides the mentioned numerous shape elements in Schaeffer's seminal work, we have indirect behavioural evidence of shape cognition in our own work:

- *Air-instruments* (Godøy et al., 2006)—studying listeners' imitations of sound-producing body motion.
- *Sound-tracing* (Nymoen et al., 2013)—studying listeners' manual shape tracings of various sound objects.

Given these instances of shapes, it seems reasonable to regard shape cognition as a mediating element in multimodality, enabling mapping from, for example, sound-producing motion shapes to auditory imagery. The very basic role of shapes in human cognition and behaviour is a fundamental element of *morpho-dynamical theory*, based on the idea that shapes are inherently holistic and efficient means for grasping and understanding complex phenomena (Thom, 1983; Petitot, 1990).

Importantly, shapes are instantaneous, that is, "all-at-once," as opposed to that which is sequential, that is, which requires a step-by-step perceptual process. We thus hypothesize that shape cognition could be a high-level code for the musculoskeletal system at the chunk timescale, in line with the theory of action gestalts (Klapp & Jagacinski, 2011), and in turn, that these action gestalts can be triggered by intermittent motor control.

Timescales

Musical imagery may involve different timescales ranging from the here-and-now of short-term memory to the long-term imagery of large-scale works, compared with the timescales discussed in Snyder (2000). We may assume that there are variable degrees of acuity in such mental images, for example, that they may be vague recollections of overall "sound" or "mood" of large-scale works, or they may be salient images of particular details. It could be useful then to distinguish the following timescales and their corresponding salient sound and motion features in view of imagery:

- *Micro* timescale with continuous features such as steady-state pitch, dynamics, and timbre, that is, as more piecewise stationary features of sound-motion objects.
- *Meso* timescale at the chunk or object timescale, that is, at the very approximate 0.5–5 seconds timescale, with salient features of dynamic, pitch-related, and timbre-related envelopes, as well as salient stylistic and aesthetic-affective features. This is also the timescale that has (variably so) been associated with sensations of the present moment (e.g., Pöppel, 1997; Varela, 1999).
- *Macro* timescale with concatenations of meso timescale chunks into large-scale musical works.

Of these three timescale categories, major constraints of music-making and music perception converge in making the meso timescale privileged in volitional musical imagery:

- Instrumental envelope constraints, that is, shapes of the excitation, sustain, and decay of sound objects, are typically found in the range of 0.5–5 seconds (Godøy, 2013).
- Coarticulation constraints, that is, that there always is a contextual smearing of both motion and sound, are likewise found in this duration range (Godøy, 2014).
- Motor control constraints necessitating hierarchical control schemes privilege the meso timescale (Grafton & Hamilton, 2007; Klapp & Jagacinski, 2011).
- Perceptual duration thresholds, defined as the minimum duration required for any feature to be manifested, for example, one measure of a waltz in order to perceive a waltz, are also typically found at the meso timescale.

We may assume that these constraints represent a *quantal element in music* (Godøy, 2013), facilitating a holistic, "all-at-once" presence of sound-motion objects in musical imagery.

Intermittency

The relationships between the continuous streams of sensory experience and the more discontinuous entities in our experience was one of the main topics of phenomenology and early gestalt theory. Husserl's famous tripartite model of *retention*, *primary impressions*, and *protention*, first published in 1893, suggested that perception and cognition proceed by a series of "now-points," by intermittently stepping out of the sensory stream in order to make sense of it as more holistic chunks (Husserl, 1928/1991). This remarkable insight could now be supplemented with ideas from recent cognitive science research. In particular, the idea of intermittency in motor control, firstly seen as a solution to various constraints, for example, the PRP, latency of body response, noise (i.e., disturbances of the neuronal information flow from the brain to the effectors), and need for anticipatory motion as in coarticulation, could now also be seen as relevant for the emergence of sound-motion objects in imagery.

In addition to theories of intermittent control (Karniel, 2013; Loram et al., 2014; Sakaguchi et al., 2015), we have corroborating elements in support of intermittency:

- detecting discontinuities in motion data (Sakaguchi et al., 2015)
- detecting muscle contractions and relaxations in EMG data (Aoki et al., 1989)
- theories of hierarchical motor control (Grafton & Hamilton, 2007)
- action gestalt theory (Klapp & Jagacinski, 2011)
- postures at salient moments in time (Rosenbaum et al., 2007; Rosenbaum, 2017)

In more detail, intermittent control "is executed as a sequence of open-loop trajectories, that is, without modification by sensory feedback apart from the instances of intermittent feedback" (Loram et al., 2014, p. 119). And furthermore:

> Using new theoretical development, methodology, and new evidence of refractoriness during sustained control, we propose that intermittent control, which incorporates a serial ballistic process within a slow feedback loop, provides the main regulation of motor effort, supplemented by fast, lower-level, continuous feedback. Refractoriness distinguishes the slow intentional from the fast reflexive loop. IC in which optimization and selection occur within the feedback loop provides powerful advantages for performance and survival. A potential neurophysiological basis for IC lies in centralized selection and optimization pathways including, respectively, the basal ganglia and cerebellum.
>
> (Loram et al., 2014, p. 124)

In short, intermittent control is a point-by-point, and top-down, motor control scheme that could be optimal both in musical performance and in musical imagery, because it entails the following:

- hierarchical and anticipatory control, bypassing the PRP and other bottlenecks;
- feedforward or so-called *open loop*, with no need for online correction;
- effector optimization by coarticulation, that is, that effectors are in place prior to use; and
- the minimization of physical effort by conservation of momentum and rebounds.

Interestingly, it has also been hypothesized that the very fast triggering by the so-called *startle effect*, that is, reaction to the sudden playback of a loud sound, may rely on a subcortical store of readymade motor programmes (Valls-Solé et al., 1999). And on a more informal note, the very fast and non-hesitant triggering of sound-producing motion in musical improvisation could likewise be possible because of such readymade motor programmes that only need an intermittent trigger to get going.

Control and Effort

For sound-producing body motion, it makes sense to see motor control and physical effort as inseparable; however, in much motor control literature, there seems to be little or no mention of physical effort. There seem to be assumptions of some kind of servomechanism, so that once a control signal has been given, the musculoskeletal system just does the rest, that is, provides the needed power for the ensuing body motion.

But in terms of imagery, it seems reasonable to claim that we not only have visual images of motion but also have salient sensations of effort, that is, of muscle activations (Jeannerod, 2001). With sensations of muscle activations as an integral element of imagery, we have come closer to establishing efficient and robust triggers for musical imagery. Brain-imaging studies have detected close relationships between various body motion tasks and the motor control areas of the brain; however, the relationships between motor control and effort, both mental and physical, seem not so well studied. Our own ongoing pupillometry studies of mental effort seem to suggest correspondences between levels of difficulty and mental effort, but we are also looking into effort issues as follows:

- motion signals, in particular velocity, acceleration, and jerk, assuming that such higher derivatives are costly in terms of effort; and
- EMG signals, detecting effort peaks and relaxation phases, including the *premotion silent period*, a relaxation phase before an effort peak (Aoki et al., 1989).

The bottom line in the motormimetic view of music perception and imagery is that all sound events, hence, all sound features, are included in some kind of body motion trajectory, and that there often will be some sense of effort associated with sounds. In view of the various constraints presented earlier in this chapter, we may hypothesize that effort is unequally distributed in time, that is, that effort is concentrated in peaks. This means that we may observe shifts between effort phases and relaxation phases in sound-producing motion. Such shifts between effort and relaxation may be regarded as advantageous because:

- For the sound-producing effectors, the concentration of effort in intermittent peaks interleaved with phases of relaxation may help avoiding strain injury in performance.
- Biomechanically, effort-relaxation shifts may help in so-called conservation of momentum by letting an already started motion continue without trying to stop it (e.g., as in bowing) or in exploiting rebounds (e.g., as in drumming).
- Shifts between effort and relaxation are also evident in coarticulation with sequences of different muscles contracting and relaxing to output the optimal sound-producing motion.

In summary, it seems that intermittency of both control and effort concerns optimization of motion, both in the real world and in mental imagery.

Triggering Imagery

For lack of more knowledge, a possible hypothesis here could be that the triggering of effort and control peaks also entails physiological sensations, similar to the mentioned startle that sets effectors in motion (Valls-Solé et al., 1999). In musical imagery, such triggers could be generated internally, but we need much more knowledge of how such triggering works. Intuitively, we may guess that there has to be some kind of threshold between non-action and action, that is, that there could be a preparatory buildup to some threshold and then a sudden triggering, that is, as a kind of capacitor model of a charge release.

Intermittent motor control in both imagery and real-world conditions could then be understood as a combination of anticipatory motor cognition, represented as sound-motion shape images, with the activation of these shape images by some intermittent trigger, in other words that we have the following:

- Open-loop, feedforward, serial ballistic triggering of sound-motion chunks with no correction once the object has been initiated, but with monitoring of the outcomes and with possibilities of correcting the features of the chunk the next time it is triggered. Thus, there are two timescales at work here, one for the intermittent triggering of motion chunks (typically in the ≈ 500 ms range), and another, much faster and more fine-grained, for monitoring the output, for example, fluctuations of pitch, dynamics, timbre, and timing in the course of the chunks' unfolding (Loram et al., 2014).
- The sound-motion chunks as coherent chunks in the sense of being singular action gestalts (Klapp & Jagacinski, 2011), that is, coherent due to the PRP at this timescale (i.e., ≈ 500 ms), evident in the performance of ornaments and other similar fast figures (Godøy et al., 2017).

Actually, we focus on various such rapid sound-motion objects, that is, mostly ornaments on various instruments, in our present work in the hope of finding out more about the triggering of anticipatory motor programmes.

Volitional Imagery for Musical Creativity

The hypothesis here is that efficient and robust volitional musical imagery may be enabled by intermittent serial ballistic triggering of sound-motion objects. This is all based on the convergence of behavioural features, first of all the phenomenon of anticipatory cognition in motor control and manifestations of this in observable body motion features (in PRP and response times, postures, trajectories, coarticulation, motion derivatives, and EMG data), and secondly on resultant musical sound features (in objects, shapes, and various patterns). This means that we may have imagery for sound-motion objects, imagery that we may test out in musical creation, first of all in improvisation and composition, but also in interpretation of scores (the transition from discrete tones to coarticulated chunks in Western music) and in mental practice for performance. An object-focused strategy as used in Schaeffer-inspired music theory teaching could be useful here, with the following categories (see Godøy, 2010, p. 60, figures 1 and 2, for examples of these object types):

- *Composed objects*, meaning the generation of incrementally different sound-motion objects by the superposition of sounds, for example, proceeding from "slim" to more "thick" objects.

- *Composite objects*, that is, making different extended objects with sounds fused by coarticulation, starting with singular sound-motion objects that then are moved so close together that they fuse and have continuous transitions between them.

More large-scale contexts are possible with concatenations of several intermittent and serial ballistic triggered objects, resulting in an emergent continuity, which in turn could be the content of instantaneous overviews of large forms “in-a-flash” as attested to by several composers, for example, Hindemith (1952/2000) and Xenakis (1992).

In summary, musical imagery may work in improvisation, composition, sound design, and so forth by a series of chunks, each generated by imagery of singular ballistic motion, each being monitored but not corrected during the unfolding, but each being corrected in the subsequent ballistic motion, as suggested by the aforementioned principles of intermittent motor control.

Conclusion

The line of reasoning here can be summarized as follows: Volitional musical imagery may be triggered by volitional motor cognition. The proximity of mental imagery and motor cognition has been quite well documented in brain-imaging studies (e.g., Lotze, 2013), but what seems less focused on is the actual triggering of body motion, both in real-world motion and in motion imagery (see also Kim, this volume). However, the idea of intermittent control does give us some novel perspectives on the temporal features of body motion and resultant musical sound, perspectives that can help us designing models to be tested on concrete music-related sound and motion data.

Intermittent motor control actually concerns the fundamental relationship between the continuous and the discontinuous in musical experience and, hence, touches on some important epistemological issues. We are now in a position to proceed from the remarkable introspective insights of Husserl on “now-points” in perception and cognition to more behavioural evidence for fundamental intermittency in perception and cognition, and in particular, to advance our knowledge and practical skills in using intermittent motor control in musical imagery.

Also, ideas presented in this chapter are closely related to artistic practice, so that the use of intermittent motor control in practical situations will be part of the further development in parallel with more conventional research. Improvisation, both in actual online live performance and in more offline compositional processes, could be a prime testing and development ground for the intermittency hypothesis.

References

Aoki, H., Tsukahara, R., & Yabe, K. (1989). Effects of pre-motion electromyographic silent period on dynamic force exertion during a rapid ballistic movement in man. *European Journal of Applied Physiology*, *58*(4), 426–432. https://doi.org/10.1007/bf00643520

Cotter, K. N. (2019). Mental control in musical imagery: A dual component model. *Frontiers in Psychology*, *10*, 1904. https://doi.org/10.3389/fpsyg.2019.01904

Gaver, W. W. (1993). What in the world do we hear? An ecological approach to auditory event perception. *Ecological Psychology*, *5*(1), 1–29. https://doi.org/10.1207/s15326969eco0501_1

Godøy, R. I. (2001). Imagined action, excitation, and resonance. In R. I. Godøy & H. Jørgensen (Eds.), *Musical imagery* (pp. 239–252). Lisse: Swets & Zeitlinger.

Godøy, R. I. (2003). Motor-mimetic Music Cognition. *Leonardo*, *36*(4), 317–319. https://doi.org/10.1162/002409403322258781

Godøy, R. I. (2006). Gestural-sonorous objects: Embodied extensions of Schaeffer's conceptual apparatus. *Organised Sound*, *11*(2), 149–157. https://doi.org/10.1017/s1355771806001439

Godøy, R. I. (2010). Images of sonic objects. *Organised Sound*, *15*(1), 54–62. https://doi.org/10.1017/s1355771809990264

Godøy, R. I. (2013). Quantal elements in musical experience. In R. Bader (Ed.), *Sound—perception—performance: Current research in systematic musicology* (Vol. 1, pp. 113–128). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-319-00107-4_4

Godøy, R. I. (2014). Understanding coarticulation in musical experience. In M. Aramaki, M. Derrien, R. Kronland-Martinet, & S. Ystad (Eds.), *Sound, music, and motion: Lecture notes in computer science* (pp. 535–547). Berlin: Springer. https://doi.org/10.1007/978-3-319-12976-1_32

Godøy, R. I. (2019). Thinking sound-motion objects. In M. Filimowicz (Ed.), *Foundations in sound design for interactive media* (pp. 161–178). New York: Routledge. https://doi.org/10.4324/9781315106342-8

Godøy, R. I., Haga, E., & Jensenius, A. (2006). Playing "air instruments": Mimicry of sound-producing gestures by novices and experts. In S. Gibet, N. Courty, & J.-F. Kamp (Eds.), *GW2005: Lecture notes in computer science* (Vol. 3881, pp. 256–267). Berlin, Heidelberg: Springer-Verlag. https://doi.org/10.1007/11678816_29

Godøy, R. I., & Leman, M. (Eds.). (2010). *Musical gestures: Sound, movement, and meaning*. New York: Routledge. https://doi.org/10.4324/9780203863411

Godøy, R. I., Song, M.-H., & Dahl, S. (2017). Exploring sound-motion textures in drum set performance. In *Proceedings of the SMC Conferences* (pp. 145–152). Espoo: Aalto University.

Gonzalez Sanchez, V. E., Dahl, S., Hatfield, J. L., & Godøy, R. I. (2019). Characterizing movement fluency in musical performance: Toward a generic measure for technology enhanced learning. *Frontiers in Psychology*, *10*, 84. doi: 10.3389/fpsyg.2019.00084.

Grafton, S. T., & Hamilton, A. F. (2007). Evidence for a distributed hierarchy of action representation in the brain. *Human Movement Science*, *26*(4), 590–616. https://doi.org/10.1016/j.humov.2007.05.009

Hindemith, P. (2000). *A composer's world: Horizons and limitations*. Mainz: Schott. (Original work published 1952)

Husserl, E. (1991). *On the phenomenology of the consciousness of internal time (1893–1917)* (J. B. Brough, Trans.; Collected Works IV). Dordrecht: Kluwer Academic Publishers. (Original work published 1928)

Jeannerod, M. (2001). Neural simulation of action: A unifying mechanism for motor cognition. *NeuroImage*, *14*(1), 103–109. https://doi.org/10.1006/nimg.2001.0832

Karniel, A. (2013). The minimum transition hypothesis for intermittent hierarchical motor control. *Frontiers in Computational Neuroscience*, *7*(12). https://doi.org/10.3389/fncom.2013.00012

Klapp, S. T., & Jagacinski, R. J. (2011). Gestalt principles in the control of motor action. *Psychological Bulletin*, *137*(3), 443–462. https://doi.org/10.1037/a0022361

Loram, I. D., van De Kamp, C., Lakie, M., Gollee, H., & Gawthrop, P. J. (2014). Does the motor system need intermittent control? *Exercise and Sport Science Review*, *42*(3), 117–125. https://doi.org/10.1249/jes.0000000000000018

Lotze, M. (2013). Kinesthetic imagery of musical performance. *Frontiers in Human Neuroscience*, *7*, 280. https://doi.org/10.3389/fnhum.2013.00280

Nanay, B. (2018). Multimodal mental imagery. *Cortex*, *105*, 125–134. https://doi.org/10.1016/j.cortex.2017.07.006

Nymoen, K., Godøy, R. I., Jensenius, A. R., & Torresen, J. (2013). Analyzing correspondence between sound objects and body motion. *ACM Transactions on Applied Perception (TAP)*, *10*(2), 1–22. https://doi.org/10.1145/2465780.2465783

Petitot, J. (1990). Forme. In *Encyclopædia Universalis*. Paris: Encyclopædia Universalis.

Pöppel, E. (1997). A hierarchical model of time perception. *Trends in Cognitive Sciences*, *1*(2), 56–61. https://doi.org/10.1016/s1364-6613(97)01008-5

Rosenbaum, D. A. (2017). *Knowing hands: The cognitive psychology of manual control*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316148525

Rosenbaum, D. A., Cohen, R. G., Jax, S. A., Weiss, D. J., & van der Wel, R. (2007). The problem of serial order in behavior: Lashley's legacy. *Human Movement Science*, *26*(4), 525–554. https://doi.org/10.1016/j.humov.2007.04.001

Sakaguchi, Y., Tanaka, M., & Inoue, Y. (2015). Adaptive intermittent control: A computational model explaining motor intermittency observed in human behavior. *Neural Networks*, *67*, 92–109. https://doi.org/10.1016/j.neunet.2015.03.012

Schaeffer, P. (1966). *Traité des objets musicaux* [Treatise on musical objects]. Paris: Éditions du Seuil.

Schaeffer, P. (with sound examples by Reibel, G., & Ferreyra, B.). (1998). *Solfège de l'objet sonore* [Music theory of the sound object]. Paris: INA/GRM. (Original work published in 1967)

Snyder, B. (2000). *Music and Memory: An Introduction*. Cambridge, MA: The MIT Press.

Thom, R. (1983). *Paraboles et catastrophes* [Parables and disasters]. Paris: Flammarion.

Valls-Solé, J., Rothwell, J. C., Goulart, F., Cossu, G., & Munoz, E. (1999). Patterned ballistic movements triggered by a startle in healthy humans. *Journal of Physiology*, *516*(3), 931–938. https://doi.org/10.1111/j.1469-7793.1999.0931u.x

Varela, F. J. (1999). The specious present: A neurophenomenology of time consciousness. In J. P. Francisco, J. Varela, B. Pachoud, & J.-M. Roy (Eds.), *Naturalizing phenomenology: Issues in contemporary phenomenology and cognitive science* (pp. 266–314). Stanford: Stanford University Press.

Williams, T. I. (2015). The classification of involuntary musical imagery: The case for earworms. *Psychomusicology: Music, Mind, and Brain*, *25*(1), 5–13. https://doi.org/10.1037/pmu0000082

Xenakis, I. (1992). *Formalized music* (Revised ed.). Stuyvesant: Pendragon Press.

4 Kinaesthetic Musical Imagery Underlying Music Cognition

Jin Hyun Kim

The various modalities of mental imagery that underlie cognition have been the object of more and more scholarly attention for quite some time. Visual modality, which has been conceived of as a dominant modality of imagistic representation, is based on Kosslyn's pictorial account for mental imagery (Kosslyn, 1980). The focus of current research on mental imagery has been either on the interplay of various modalities (Schaefer, 2014) or on a modality of mental imagery that is different from a modality of sensory stimulation (cf. Nanay, this volume, for an overview). It is in this context that kinaesthetic modality has recently begun to be explored more deeply. While, in general, kinaesthetic imagery can be described as mental imagery involving movement sense, the bulk of the research on the kinaesthetic modality of mental imagery has been on motor imagery (Hanakawa, 2016, Jeannerod, 1994; Malouin et al., 2007; McNeill et al., 2020) that is related to mental simulation of the execution of an agent's action (cf. Decety & Grèzes, 2006). In this research context, kinaesthetic imagery is conceived of as imagery that involves the body sense associated with movement from one's own perspective rather than involving the perspective of an external observer (Lorey et al., 2009). Music research, however, usually uses the terms *kinaesthetic imagery* and *motor imagery* (Lotze, 2013) interchangeably, or describes kinaesthetic components of mental imagery underlying music cognition when discussing the role of motor imagery, which is considered to underlie the imagining (rehearsing, mentally training, and mentally imitating) actions required for producing musical sounds and music-making (Bailes, 2019; Brown & Palmer, 2013; Cox, 2011, 2016; Godøy, 2001, 2010). Against this background, this chapter argues that *kinaesthetic musical imagery* can be understood as a category of mental imagery that could be best characterized as the (quasi-)perceptual conscious experience of dynamic self-movement that mental representations of musical dynamic properties give rise to (Part 4), drawing on research on mental imagery (Part 1), kinaesthesia (Part 2), and phenomenal consciousness (Part 4). Since this chapter aims to distinguish kinaesthetic imagery that does not involve mental simulation of motor action from motor imagery, one section (Part 3) is devoted to delineating kinaesthetic imagery in opposition to motor imagery, followed by a positive definition of kinaesthetic (musical) imagery (Part 4).

DOI: 10.4324/9780429330070-6

Part 1: Mental Imagery

Mental imagery can be defined as a basic mode of representation. Although there has been a debate about whether this mode is propositional, supported by Pylyshyn (1973), or pictorial, initiated by Kosslyn (1980), mental imagery has been mostly conceived of as an imagistic mode rather than a propositional one (Paivio, 1986; van Eckhardt, 1999). Taking into account how the representational content is given, Eysenck and Keane (1995) state that propositional representations are amodal whereas imagistic representations are modality-specific. A non-propositional mode of representation accounts for a concept of representation that addresses not only high-ordered cognition (such as problem-solving, reasoning, and thinking) but also lower-level cognition (such as processes involved in motor control), which does not involve a language of thought and plays an important role during music-making and music perception.

The term *mental imagery* can refer to (quasi-)perceptual conscious experience per se or modality-specific mental representations that directly call forth (quasi-) perceptual conscious experience (Thomas, 2019). Mental imagery can then be described as quasi-perceptual conscious experience if it occurs in the absence of an external stimulus, yet its structure in our consciousness is analogous to that of a percept. This definition allows to subsume different definitions of musical imagery: first, as the experience of imagining musical sounds in the absence of corresponding sensory stimulation (cf. Bailes, 2007, p. 555; Godøy & Jørgensen, 2001, p. ix) related to a form of modality-relevant offline cognition (Bailes, 2019, p. 447) and, secondly, as the cognitive process underlying music-making and music perception (cf. Juslin, 2013; Juslin & Västfjäll, 2008; Taruffi & Küssner, 2019).

Propositional representations are discrete and explicit, and involve strong combination rules (Eysenck & Keane, 1995). In contrast, imagistic representations are non-discrete and implicit, and involve loose combination rules. Accordingly, the concept of mental imagery allows to rethink not only the modality of representation related to content but also the form of representation (as to whether it is discrete or indiscrete, explicit or implicit, etc.; van Eckhardt, 1999).

Part 2: A Kinaesthetic Modality of Mental Imagery

The term *kinaesthesia*, referring to the ability to sense movement, was coined by Henry Charlton Bastian in his *The Brain as an Organ of Mind* (1880). Kinaesthesia is a specific form of proprioception pertaining only to one's own movement (Rosenbaum, 1991, p. 51). While proprioception is anchored in external sensory receptors within the muscles, tendons, joints, and skin that register felt bodily changes and movement, kinaesthetic information is provided by internally mediated neuro-muscular systems that are not dependent on external stimuli (Sheets-Johnstone, 2019, p. 144f.). Kinaesthesia is involved both in the execution of bodily movement and in perception via any modality (visual, auditory, etc.; Berthoz & Petit, 2006/2008). Since kinaesthesia is involved in every process of sensory perception, it could be conjectured that the kinaesthetic modality of mental imagery forms the basis for all imagistic representation.

Indeed, the kinaesthetic modality of mental imagery is not necessarily triggered by the perception of kinaesthetic stimulation during the execution of bodily movement. Instead, kinaesthetic imagery can represent the dynamic structure of movements shaped during execution. This distinction between kinaesthetic imagery and the perception of kinaesthetic stimulation is similar to the American psychologist Edward Bradford Titchener's distinction between kinaesthetic image and kinaesthetic sensation. Titchener discussed this distinction when he introduced the term *empathy* as a translation of the German term *Einfühlung* that Theodor Lipps coined in the context of psychological aesthetics in the late 19th century. Titchener uses the term *kinaesthetic image* to account for Lipps's notion of inner experience underlying aesthetic empathy. Lipps suggests that there is a process underlying the empathic aesthetic experience of art, which he terms an *inner experienced doing* (Lipps, 1903, p. 187, my translation). He characterizes this process as a kind of image of movement, distinguishing it from organic sensations such as muscle contraction or relaxation (Lipps, 1903). According to Titchener, kinaesthetic image is concerned with feeling movement properties "in the mind's muscle" (Titchener, 1909, p. 21), which appears as an appropriate characterization for the modern term *kinaesthetic imagery*. It is worth noting that some movement properties that Titchener lists as examples are gravity (not in the sense of the physical attraction of bodies, but the emotional depth or solemnity of experience), modesty, pride, courtesy, and stateliness (Titchener, 1909, p. 21)—not the kinds of activities that could be *executed* with the mind's muscle.

In the context of the topic of this chapter, *kinaesthetic musical imagery*, and keeping in mind Lipps' and Titchener's notion of an image of movement or kinaesthetic image, it seems to be helpful to discuss what kinaesthetic imagery is *not*, by briefly reviewing relevant research related to musical imagery.

Part 3: What Kinaesthetic Musical Imagery Is Not

First, kinaesthetic musical imagery is not the same as imagery involving the simulation of motor actions required for sound generation and music-making (Cox, 2016; Godøy, 2001). This simulation correlates with the brain's supplementary motor area (SMA) activities and is also at play during the execution and perception of the same action. In cognition research in general and music research in particular, imagery involving the simulation of motor actions is known as *motor imagery*. Music research also uses motor imagery for mental training and rehearsal of playing musical instruments (Brown & Palmer, 2013; for an overview, see Bailes, 2019; Lotze, 2013). The motor representations employed for actual movements form the basis for motor imagery (Lotze, 2013, p. 2). Lotze (2013, pp. 3ff.) argues that experience in actual movement performance is a prerequisite for the activity of motor-related areas of the brain, such as supplementary motor area (SMA), the premotor cortex (PMC), parietal areas, and the cerebellum. These areas are assumed to be responsible for motor imagery. Godøy

(2001, 2010) discusses a close link between mental imagery of sound and sound-producing bodily action. He also argues that mental imagery of sound could be triggered by that of sound-producing bodily action (Godøy, 2010; see Godøy, this volume). In line with Godøy's proposition, Cox (2011, 2016) develops principles of what he calls mimetic comprehension of music. His first two principles are closely related to Godøy's approach: "1. Sounds are produced by physical events: Sounds indicate (signify) the physicality of their source; 2. Many or most musical sounds are evidence of the *human* motor actions that produce them" (Cox, 2016, p. 13). He then suggests that mimetic behaviour, that is, behaviour imitating the observed actions of others, be it overt or covert, plays a significant role in human understanding of other entities. Covert aspects of mimetic behaviour, which he refers to as *mimetic motor imagery*, are particularly relevant for the topic of the present chapter. For Cox (2016, p. 12), motor imagery is a mental representation, and mimetic motor imagery is a representation of observed actions, which he characterizes as mimetic (Cox, 2016, p. 15) and bodily (Cox, 2016, p. 13). The notion of bodily representation is related to the recruitment of brain areas representing motor actions during imagery (Cox, 2016, p. 15). According to Cox and Godøy, mental imagery underlying the act of listening to music is imagery that is mentally re-enacting actions that are necessary to produce sounds and shape musical features such as pitch and timbre. Cox (2016, p. 14) characterizes this mental re-enactment of motor actions as mimetic representation.

On the other hand, the study by Eitan and Granot (2006) that investigates listeners' images of motion reveals another aspect under which imagery related to motion, which they term *spatial and kinetic imagery* (Eitan & Granot, 2006, p. 222), does not necessarily represent sound-producing bodily action. Eitan and Granot are interested in how listeners associate changes in musical parameters with physical space and bodily motion. In their study, bodily motion is limited to imagined movements of "a *human* character" (Eitan & Granot, 2006, p. 221; my emphasis). This study focuses on associations of the type, direction, and pace change of these motions and on the forces affecting them with intensification and abatement of musical parameters, including dynamics, pitch contour, pitch intervals, attack rate, and articulation. Results show that there is a strong link between change in musical parameters and spatial and kinaesthetic imagery. Musical intensifications are linked to increasing speed, and musical abatements are closely linked to spatial descents.

However, consideration of other studies calls for a critical examination of the claim that mental imagery of sound is either imagery associated with motions of a *human* character in general or simulation of sound-producing bodily action in particular. For example, Todd (1992) claims that the music's expressive properties induce a sense of self-movement, unlike a sense of an external agent in motion as assumed even in the study by Eitan and Granot (2006). Gjerdingen's (1994) notion that the experience of music is associated with motion in virtual space in a metaphorical sense provides another example.

In a study on the mental imagery at play while listening to music that was conducted with musicology students at Humboldt-Universität zu Berlin in 2019, I carried out micro-phenomenological interviews. Micro-phenomenological interviews allow the interviewer to take a second-person perspective to help the interviewee re-live their previous first-person experience and to access the dynamic process of that experience (cf. Petitmengin, 2006; Petitmengin et al., 2018). Two persons who heard the same piece of music (Symphony No. 4 in C minor *War and Peace* by Joachim Nicolas Eggert, 1812) reported that one specific musical passage evoked in them images of water falling, and they became aware of their own mental movement during the interviews—in one case, of mentally sauntering, and in the other case, of mentally swaying. The latter interviewee said: “There is a passage where I see water falling motions mentally, in visual imagery. At the same time, I feel a kind of motion. Yes, I am swaying mentally.” In another interview study using a micro-phenomenological interview technique on the experience of musicians empathizing with their own performance, one pianist reported that whenever she feels at one with her performance, she sees an abstract line that moves along her performance. When I asked what it feels like to see an abstract line moving, she answered that the abstract line that moves is what she *feels* mentally, not something she sees in her mind’s eye. Asked again whether she does not see that line mentally, she confirmed that she traces her own mental movement that feels like the movement of a line, but she does not mentally see a line moving. This indicates that she experiences kinaesthetic imagery in these moments, but without visual imagery. This kind of kinaesthetic imagery pertains to self-movement rather than simulation of the actions of others.

Secondly, kinaesthetic imagery is not the same as the imagination of (imitated) motor action, be it sound-producing and music-shaping or not. According to Cox (2016, p. 12), even mimetic motor imagery is not always the imagination of (imitated) motor action, since imagination is a conscious and deliberate act, whereas imagery includes an involuntary and automatic process. In cases where mimetic motor imagery is the imagination of (imitated) motor action, imagination could be related to motor action that a person has already performed, such as a sound-producing bodily action while playing a musical instrument. Alternatively, it could be about imitated motor action. One can think in two ways about imitated motor action: one could imagine motor actions that have imitatively been executed through observing a musical performance or listening to music, relating such motor actions to sound-producing bodily actions. Cox (2011, 2016) refers to this sort of imitated motor action as *mimetic motor action*. One could also imagine other sorts of motor actions imitating musical properties. The imagination of imitated motor action is neither mimetic motor action nor any kind of overt action, but covert action. Moreover, this kind of imagination is meta-representational, as imitated motor action is itself a representation that is directed towards an event in the world, and imagination is yet another representation that is directed towards imitated motor action.

Kinaesthetic imagery, which could underlie the imagination of motion, however, is neither necessarily the imagination of previously performed motor action nor the imagination of imitated motor action that could be executed during perception. While mimetic motor imagery involves meta-representational processes, kinaesthetic imagery that is not any imagination of (imitated) motor action is about awareness of self-movement that can be overt or covert. Although kinaesthetic imagery is not based on meta-representational processes, it can be considered to involve understanding processes, since it can imitatively single out relevant musical properties, such as rhythm and dynamics, in a way that takes a specific perspective and shows a structural correspondence between imagery and perceived musical properties. During the micro-phenomenological interviews within the scope of the aforementioned study on mental imagery, several subjects reported such experiences: "I become aware of my inner movement whenever I notice that musical rhythm gets more complex than before"; "I am tracing melancholy through the whole piece. Moreover, I am moving internally, feeling tension when the piece triggers the climax' feeling, a peak, accompanied by particular tones, and released after that moment." Since this kind of imagery is intentionally directed towards a perceived object, involves the conditions of satisfaction, and generates a structural correspondence between imagery and perceived object, it could be conceived of as representation (albeit not for high-ordered cognition).

Part 4: Phenomenal Consciousness and Kinaesthetic (Musical) Imagery

Kinaesthetic imagery can be best explained within the framework of theories of consciousness. Phenomenal consciousness is about mental states in which experiential qualities are manifested. It allows the individual to trace what it is like to be in a certain mental state (Block, 1995, p. 230). For instance, when we hear musical sound events, we hear them in such a specifically inward way that we can trace what it is like to hear each event, the transitions between those events, and what it is like to feel those dynamics. Phenomenal consciousness is directed towards those experiential properties, not towards a series of musical events—although there could be a correspondence between musical properties and experiential properties.

According to Damasio (2010), phenomenal consciousness depends on self-consciousness. The foundation of what he terms *core consciousness* (Damasio, 1999) is a preconscious precedent that is referred to as *proto-self*, which is "a specific set of neural structures [that] support[s] the first-order representation of current body states" (Damasio, 1999, p. 159). Changes in neural activity patterns and body states emerge while perceiving an object, responding to sensory stimuli, and interacting in an environment. When we become aware of changes in the body state, "low-level attention" (Damasio, 1999, p. 89), "mental streams of images" (Damasio, 1999, p. 88), and "background emotion" (Damasio, 1999, p. 89) emerge. This level of awareness is referred to as *core consciousness*. A mental state that comes to the foreground in core consciousness is not merely a reactive state directed towards sensory stimulation. Rather, it is a feeling of the sensing of body states, which is directed towards "an internally-mediated . . . awareness" (Sheets-Johnstone, 2019, p. 145) such as the sensing of "a particular flow of movement" (Sheets-Johnstone, 2019, p. 161), and hence belongs to phenomenal consciousness. It is at this layer of consciousness that mental imagery emerges, since mental imagery is a mode of representation that is caused by neural activity patterns as well as the (quasi-)perceptual conscious experience that mental representations give rise to.

Kinaesthetic imagery represents the dynamic properties of a perceived object and gives rise to the experience of dynamic self-movement, that is, the experience that I am the subject of the movement. The dynamic properties of a perceived object do not necessarily involve a sense of an external agent. It can be related to properties of an inanimate object that are associated with dynamic properties of animate movement that allow, for instance, the experience of a series of tension and relaxation. Moreover, kinaesthetic imagery could be related to the interoception of physiological processes, such as a change in heartbeat rate and respiratory rate that is entrained to dynamic change in sensory stimuli. Although it is not clear yet whether interoceptive processes shape our mental imagery (Bailes, 2019, p. 458), conscious interoception could be considered to trace what it is like to be exposed to those physiological processes, and therefore to belong to phenomenal consciousness. Kinaesthetic imagery also serves as a basis for the experience of dynamic quality of the self, especially in the context of aesthetic experience. This is similar to Lipps's *inner experienced doing* as a process underlying empathic aesthetic experience, which can be interpreted as kinaesthetic imagery, as was briefly discussed in Part 2. According to Lipps's theory of aesthetic empathy, phenomenal consciousness of self is experienced due to such inner activity, which accompanies the perception and appreciation of an aesthetic object. In other words, one feels oneself as being active during the aesthetic experience (Lipps, 1900).

Given the aforementioned, *kinaesthetic musical imagery* could be defined as the (quasi-)perceptual conscious experience of dynamic self-movement that mental representations of musical dynamic properties give rise to. Musical dynamic properties can be described as what the developmental psychologist Daniel Stern terms *forms of vitality*, which allow the experience of being alive, creating a sense of movement, time, space, force, and directionality in the human mind (Stern, 2010, p. 4). This means that forms of vitality manifest in structural features (e.g., contours of dynamics and pitch) and codified forms (e.g., rondo form) (Kim, 2013; Stern, 2010). Forms of vitality could manifest themselves even in neural and bodily activities affected by musical dynamic properties, allowing for entrainment of the rhythms of neural and bodily activities to musical rhythm. Based on dynamic properties shared among musical structures, structures of neural and bodily activities that occur in reaction to musical stimuli, and structures of mental representations, it could be assumed that kinaesthetic musical imagery that builds from neural activity patterns and the body state, and belongs to phenomenal consciousness, relates to dynamic self-movement without any reference to an external agent.

Conclusion

Against the background of an ever-growing body of music research that is interested in the kinaesthetic modality of mental imagery underlying music cognition, this chapter focused on the concept of kinaesthetic musical imagery. Having observed that music research often uses the term *kinaesthetic imagery* as a synonym for *motor imagery*, or describes kinaesthetic components of motor imagery, which is related to mental simulation of the execution of goal-directed action, it was necessary to discuss the relationship between kinaesthetic imagery and motor imagery.

Starting from the general description of kinaesthetic imagery as imagery involving the body sense associated with movement, the question of whether kinaesthetic musical imagery involves a sense of an external agent, which is supposed in the case of motor imagery, deserved careful discussion. Taking into account empirical evidence that musical dynamic properties can induce a sense of self-movement, which could be described in terms of kinaesthetic musical imagery, I argued that kinaesthetic musical imagery is neither the same as imagery involving the simulation of motor actions required to produce musical sounds and to shape music nor the same as the imagination of (imitated) motor action, be it sound-producing and music-shaping or not. Rather, kinaesthetic musical imagery should be understood as (quasi-)perceptual conscious experience related to dynamic self-movement, which is accessible to phenomenal consciousness. This kind of experience is brought forth by mental representations of musical dynamic properties, whose very structures correspond to musical structures.

References

Bailes, F. (2007). The prevalence and nature of imagined music in the everyday lives of music students. *Psychology of Music*, *35*(4), 555–570. doi:10.1177/0305735607077834

Bailes, F. (2019). Empirical musical imagery beyond the "mind's ear". In. M. Grimshaw-Aagaard, M. Walther-Hansen, & M. Knakkergaard (Eds.), *The Oxford handbook of sound and imagination* (Vol. 2, pp. 445–463). New York: Oxford University Press.

Bastian, H. C. (1880). *The brain as an organ of mind*. London: Kegan Paul.

Berthoz, A., & Petit, J. L. (2008). *The physiology and phenomenology of action*. New York: Oxford University Press. (Original work published 2006)

Block, N. (1995). On a confusion about the function of consciousness. *Behavioral and Brain Sciences*, *18*(2), 227–247. https://doi.org/10.1017/s0140525x00038188

Brown, R. M., & Palmer, C. (2013). Auditory and motor imagery modulate learning in music performance. *Frontiers in Human Neuroscience*, *7*(320), 1–13. https://doi.org/10.3389/fnhum.2013.00320

Cox, A. (2011). Embodying music: Principles of the mimetic hypothesis. *Music Theory Online 17*(2). doi:10.30535/mto.17.2.1

Cox, A. (2016). *Music and embodied cognition: Listening, moving, feeling, and thinking*. Bloomington: Indiana University Press.

Damasio, A. (1999). *The feeling of what happens: Body and emotion in the making of consciousness*. New York: Harcourt Brace & Company.

Damasio, A. (2010). *Self comes to mind: Constructing the conscious brain*. New York: Pantheon Books.

Decety, J., & Grèzes, J. (2006). The power of simulation: Imagining one's own and other's behavior. *Brain Research*, *1079*(1), 4–14. doi: 10.1016/j.brainres.2005.12.115

Eitan, Z., & Granot, R. Y. (2006). How music moves: Musical parameters and listeners' images of motion. *Music Perception*, *23*(3), 221–247. doi: 10.1525/mp.2006.23.3.221

Eysenck, M. W., & Keane, M. T. (1995). *Cognitive psychology: A student's handbook*. Hillsdale: Erlbaum.

Gjerdingen, R. O. (1994). Apparent motion in music? *Music Perception*, *11*(4), 335–370. https://doi.org/10.2307/40285631

Godøy, R. I. (2001). Imagined action, excitation, and resonance. In R. I. Godøy & H. Jørgensen (Eds.), *Musical imagery* (pp. 237–259). Lisse: Swets & Zeitlinger.

Godøy, R. I. (2010). Images of sonic objects. *Organised Sound*, *15*(1), 54–62. https://doi.org/10.1017/s1355771809990264

Godøy, R. I., & Jørgensen, H. (Eds.). (2001). *Musical imagery*. Lisse: Swets & Zeitlinger.

Hanakawa, T. (2016). Organizing motor imageries. *Neuroscience Research*, *104*, 56–63. doi:10.1016/j.neures.2015.11.003

Jeannerod, M. (1994). The representing brain. Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences*, *17*(2), 187–202. doi:10.1017/S0140525X00034026

Juslin, P. N. (2013). From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews*, *10*(3), 235–266. doi:10.1016/j.plrev.2013.05.008

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, *31*(5), 559–575. doi:10.1017/S0140525X08005293

Kim, J. H. (2013). Shaping and co-shaping forms of vitality in music: Beyond cognitivist and emotivist approaches to musical expressiveness. *Empirical Musicology Review*, *8*(3–4), 162–173. https://doi.org/10.18061/emr.v8i3-4.3937

Kosslyn, S. M. (1980). *Image and mind*. Cambridge, MA: Harvard University Press.

Lipps, T. (1900). Aesthetische Einfühlung. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, *22*, 415–450.

Lipps, T. (1903). Einfühlung, innere Nachahmung, und Organempfindungen. *Archiv für die gesamte Psychologie*, *1*, 185–204.

Lorey, B., Bischoff, M., Pilgramm, S., Stark, R., Munzert, J., & Zentgraf, K. (2009). The embodied nature of motor imagery: The influence of posture and perspective. *Experimental Brain Research*, *194*(2), 233–243. doi:10.1007/s00221-008-1693-1

Lotze, M. (2013). Kinaesthetic imagery of musical performance. *Frontiers in Human Neuroscience*, *7*, 280. doi:10.3389/fnhum.2013.00280

Malouin, F., Richards, C. L., Jackson, P. L., Lafleur, M. F., Durand, A., & Doyon, J. (2007). The Kinesthetic and Visual Imagery Questionnaire (KVIQ) for assessing motor imagery in persons with physical disabilities: A reliability and construct validity study. *Journal of Neurologic Physical Therapy*, *31*(1), 20–29. https://doi.org/10.1097/01.npt.0000260567.24122.64

McNeill, E., Ramsbottom, N., Toth, A. J., & Campbell, M. J. (2020). Kinaesthetic imagery ability moderates the effect of an AO+MI intervention on golf putt performance: A pilot study. *Psychology of Sport & Exercise*, *46*(101610). doi:10.1016/j.psychsport.2019.101610

Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford: Oxford University Press.

Petitmengin, C. (2006). Describing one's subjective experience in the second person: An interview method for the science of consciousness. *Phenomenology and Cognitive Science*, *5*(3–4), 229–269. https://doi.org/10.1007/s11097-006-9022-2

Petitmengin, C., Remillieux, A., & Valenzuela-Moguillansky, C. (2018). Discovering the structures of lived experience: Towards a micro-phenomenological analysis method.

Phenomenology and the Cognitive Sciences, *18*(4), 691–730. https://doi.org/10.1007/s11097-018-9597-4

Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin*, *80*, 1–24. https://doi.org/10.1037/h0034650

Rosenbaum, D. A. (1991). *Human motor control*. San Diego: Academic Press.

Schaefer, R. S. (2014). Mental representations in musical processing and their role in action-perception loops. *Empirical Musicology Review*, *9*(3–4), 161–176. https://doi.org/10.18061/emr.v9i3-4.4291

Sheets-Johnstone, M. (2019). Kinesthesia: An extended critical overview and a beginning phenomenology of learning. *Continental Philosophy Review*, *52*(2), 143–169. https://doi.org/10.1007/s11007-018-09460-7

Stern, D. N. (2010). *Forms of vitality: Exploring dynamic experience in psychology, the arts, psychotherapy, and development*. New York: Oxford University Press.

Taruffi, L., & Küssner, M. (2019). A review of music-evoked visual mental imagery: Conceptual issues, relation to emotion, and functional outcome. *Psychomusicology: Music, Mind, and Brain*, *29*(2–3), 62–74. doi:10.1037/pmu0000226

Thomas, N. J. Z. (2019). Mental representation. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Summer 2019 ed.). Stanford University. https://plato.stanford.edu/archives/sum2019/entries/mental-imagery/

Titchener, E. B. (1909). *Lectures on the experimental psychology of thought-processes*. New York: The Macmillan Company.

Todd, N. P. M. (1992). The dynamics of dynamics: A model of musical expression. *Journal of the Acoustical Society of America*, *91*(6), 3540–3550. https://doi.org/10.1121/1.402843

Van Eckhardt, B. (1999). Mental representation. In R. A. Wilson & F. C. Keil (Eds.), *The MIT encyclopedia of the cognitive sciences* (pp. 527–529). Cambridge: MIT Press.

5 Music and Multimodal Mental Imagery

Bence Nanay

Mental imagery is early perceptual processing that is not triggered by corresponding sensory stimulation in the relevant sense modality. *Multimodal mental imagery* is early perceptual processing that is triggered by sensory stimulation in a different sense modality. For example, when early visual or tactile processing is triggered by auditory sensory stimulation, this amounts to multimodal mental imagery. Pulling together philosophy, psychology, and neuroscience, I will argue in this chapter that multimodal mental imagery plays a crucial role in our engagement with musical works. Engagement with musical works is normally a multimodal phenomenon, where we get input from a number of sense modalities. But even if we screen out any input that is not auditory, multimodal mental imagery will still play an important role that musicians and composers often actively rely on.

The plan of the chapter is the following. In Part 1, I outline a fairly wide conception of mental imagery, which is very close to how psychologists and neuroscientists use this concept and zero in on an important, but thus far relatively unexplored form of mental imagery, which I call multimodal mental imagery. In Part 2, I highlight the importance of auditory mental imagery in listening to music. In Part 3, I argue that multimodal mental imagery also plays a crucial role in listening to music. Finally, in Part 4, I show why we should all care about this.

Part 1: Multimodal Mental Imagery

Mental imagery, as psychologists and neuroscientists understand the concept, is early perceptual processing not triggered by corresponding sensory stimulation in the relevant sense modality (Kosslyn et al., 2006; Nanay, 2015, 2018, 2023; Pearson et al., 2015). For example, *visual imagery* is early visual processing (say, in the primary visual cortex) not triggered by corresponding sensory stimulation in the visual sense modality (that is, not triggered by corresponding retinal input).

Mental imagery is defined negatively: it is early perceptual processing *not* triggered by corresponding sensory stimulation in the relevant sense modality. But then what is it triggered by? It can be triggered, in a purely top-down manner, say, when you close your eyes and visualize an apple. But it can also be triggered laterally by other sense modalities. This is multimodal mental imagery. *Multimodal*

DOI: 10.4324/9780429330070-7

mental imagery is early perceptual processing in one sense modality triggered by sensory stimulation in a different sense modality.

When you have early auditory processing that is triggered by sensory stimulation in the visual sense modality, this counts as multimodal mental imagery. This is what happens, when, for example, you watch the television muted (see, e.g., Calvert et al., 1997; Hertrich et al., 2011; Pekkola et al., 2005). Conversely, if you have early visual processing that is triggered by sensory stimulation in another sense modality, this also counts as multimodal mental imagery.

Another example of multimodal mental imagery (which I will come back to) is the double flash illusion. You are presented with one flash and two beeps simultaneously (Shams et al., 2000). So the sensory stimulation in the visual sense modality is one flash. But you experience two flashes, and already in the primary visual cortex, two flashes are processed (Watkins et al., 2006). This means that the double flash illusion is really about multimodal mental imagery: we have perceptual processing in the visual sense modality (again, already in V1) that is not triggered by corresponding sensory stimulation in the visual sense modality (but by sensory stimulation in the auditory sense modality).

I need to make some clarificatory remarks about the concept of multimodal mental imagery. First, not everybody will agree with my use of the term "mental imagery" (which I borrow from the standard usage in psychology and neuroscience). Nothing depends on the label "mental imagery" in the argument that follows. Those who have very strong views about how this concept should or should not be used should read the rest of the argument to be about mental imagery* (which is defined as early perceptual processing not triggered by corresponding sensory stimulation in the relevant sense modality).

Secondly, mental imagery may or may not be voluntary. When you close your eyes and visualize an apple, this is an instance of voluntary mental imagery. But not all mental imagery is voluntary. Unwanted flashbacks to an unpleasant scene or earworms in the auditory sense modality would be examples for involuntary mental imagery (see Floridou, this volume).

Thirdly, mental imagery may or may not localize its object in one's egocentric space. When we visualize an apple, we often do so in such a way that the apple is represented in some kind of abstract visualized space, so that it would make little sense to ask whether you could reach the apple or how far the apple is from the tip of your nose. But this is, again, not a necessary feature of mental imagery. We can also visualize an apple on the pages of this book you are reading.

Finally, and most controversially, perceptual processing may be conscious or unconscious. We know from many experimental studies that perception, that is, sensory stimulation-driven perceptual processing, can be unconscious (Kentridge et al., 1999; Kouider & Dehaene, 2007; Weiskrantz, 2009). So there is no prima facie reason why mental imagery, that is, non-sensory stimulation-driven perceptual processing, would have to be conscious. Just as perceptual processing in general, perceptual processing that is not triggered by corresponding sensory stimulation in the relevant sense modality may also be conscious or unconscious (see Nanay, 2018, 2021).

Part 2: Music and Auditory Mental Imagery

Expectations play a crucial role in our engagement with music (and art in general). When we are listening to a song, even when we hear it for the first time, we have some expectations of how it will continue. And when it is a tune we are familiar with, this expectation can be quite strong (and easy to study experimentally). When we hear "Ta-Ta-Ta" at the beginning of the first movement of Beethoven's *Fifth Symphony* in C minor, Op. 67 (1808), we will strongly anticipate the closing "Taaam" of the "Ta-Ta-Ta-Taaaam."

Many of our expectations are fairly indeterminate: when we are listening to a musical piece we have never heard before, we will still have some expectations of how a tune will continue, but we don't know what exactly will happen. We can rule out that the violin glissando will continue with the sounds of a beeping alarm clock (unless it's a really avant-garde piece . . .), but we can't predict with great certainty how exactly it will continue. Our expectations are malleable and dynamic: they change as we listen to the piece.

Expectations are mental states that are about how the musical piece will unfold. So they are future-directed mental states. But this leaves open just what kind of mental states they are—how they are structured, how they represent this upcoming future event, and so on. I will argue that at least some forms of expectations in fact count as mental imagery. And musical expectations (of the kind involved in examples like the "Ta-Ta-Ta-Taaaam") count as auditory temporal mental imagery.

Temporal mental imagery is early perceptual processing that is not triggered by temporally corresponding sensory stimulation in the relevant sense modality. In other words, temporal mental imagery is perceptual processing that comes either too late or too early compared to the sensory stimulation. Some expectations count as temporal mental imagery whereby the perceptual processing comes before the expected sensory input (and in this sense, it is not triggered by temporally corresponding sensory input).

One example is the recently discovered time-compressed wave of activity in the primary visual cortex. This only happens if there is a familiar sequence of sensory events (whereby sensory inputs follow each other predictably and in a way that has been entrenched by past experiences). On the trigger of the beginning of this sequence, there is time-compressed activity in the primary visual cortex (Ekman et al., 2017). This is quite literally perceptual processing (and very early perceptual processing) that is not triggered by temporally corresponding sensory stimulation in the relevant sense modality. The primary visual cortex is activated in a way that would correspond to the sensory input that is most likely to come in the next milliseconds.

Another well-researched example of expectations as mental imagery comes from pain research. Pain very much depends on our expectations (Atlas & Wager, 2012; Goffaux et al., 2007; Koyama et al., 2005; see Peerdeman et al., 2016, for a meta-analysis), and there is more and more data on the neural mechanism of this process (Jensen et al., 2003; Keltner et al., 2006; Ploghaus et al., 1999).

One crucial finding is about the sensory cortical areas (S1/S2). These brain regions are very early cortical areas involved in the processing of painful input very early on in cortical processing. This happens when you get a papercut—in the case of pain driven by sensory input. But it also happens when you have expectations about getting a papercut—in the case of expectations of pain (Porro et al., 2002; Wager et al., 2004). In other words, some expectations of pain will count as mental imagery in the psychological/neuroscientific sense of the term as we have clear evidence that some expectations can activate S1/S2 without any nociceptors (pain receptors) being involved (Keltner et al., 2006; Ploghaus et al., 1999; Porro et al., 2002; Wager et al., 2004), and this amounts to early cortical pain processing that is not triggered by corresponding painful input.

It is important to emphasize that this does not mean that expectations in general will have to be labelled as mental imagery. Many instances of expectations will not count as mental imagery—for example, if I make an appointment with my dentist for next month and I am anticipating the pain I will have to endure then, this will not count as an instance of mental imagery as long as the somatosensory cortical areas are not directly activated. But we have plenty of evidence that at least some expectations can activate the somatosensory cortical areas directly, without the involvement of nociceptors (see Nanay, 2017, for a summary). These instances of expectations will count as mental imagery.

Similarly, for musical expectations: if I anticipate a dreadful experience at the concert tomorrow, this is not necessarily a form of mental imagery (although it might be accompanied by mental imagery). But the way my mind expects a certain note in the next split second is a form of auditory temporal mental imagery.

Take, again, the toy example of Beethoven's Fifth. On this view, the listener forms a mental imagery of the fourth note ("Taaam") on the basis of the experience of the first three ("Ta-Ta-Ta," there is a lot of empirical evidence that this is in fact what happens; see Herholz et al., 2012; Kraemer et al., 2005; Leaver et al., 2009; Yokosawa et al., 2013; Zatorre & Halpern, 2005). This mental imagery may or may not be conscious. But if the actual "Taaam" diverges from the way our mental imagery represents it (again, if it is delayed, or altered, e.g., in pitch or timbre), we notice this divergence and experience its salience in virtue of a noticed mismatch between the experience and the mental imagery that preceded it.

The "Ta-Ta-Ta-Taaam" example is a bit simplified, so here is a real-life and very evocative case study, an installation by the British artist, Katie Paterson. The installation is an empty room with a grand piano in it, which plays automatically. It plays a truncated version of Beethoven's Moonlight Sonata. The title of the installation is *Earth-Moon-Earth (Moonlight Sonata Reflected From the Surface of the Moon)* (2007). Earth-Moon-Earth is a form of transmission (between two locations on Earth), where Morse codes are beamed up the moon, and they are reflected back to earth. While this is an efficient way of communicating between two far-away (Earth-based) locations, some information is inevitably lost (mainly because some of the light does not get reflected back but it is absorbed in the Moon's craters).

In *Earth-Moon-Earth (Moonlight Sonata Reflected From The Surface of The Moon* (2007), the piano plays the notes that did get through the Earth-Moon-Earth transmission system, which is most of the notes, but some notes are skipped. Listening to the music the piano plays in this installation, if you know the piece, your auditory mental imagery is constantly active, filling in the gaps where the notes are skipped.

Part 3: Multimodal Perception and Multimodal Mental Imagery When Listening to Music

There is a lot of recent evidence that multimodal perception is the norm and not the exception—our sense modalities interact in a variety of ways (see Spence & Driver, 2004; Vroomen et al., 2001, for summaries). Information in one sense modality can influence and even initiate information processing in another sense modality at a very early stage of perceptual processing.

A simple example is *ventriloquism*, which is commonly described as an illusory auditory experience influenced by something visible (Bertelson, 1999). It is one of the paradigmatic cases of cross-modal illusion: We experience the voices as coming from the dummy, while they in fact come from the ventriloquist. The auditory sense modality identifies the ventriloquist as the source of the voices, while the visual sense modality identifies the dummy. And, as it often happens in cross-modal illusions, the visual sense modality wins out: our (auditory) experience is of the voices as coming from the dummy.

Further, as we have seen, if there is a flash in your visual scene and you hear two beeps during the flash, you experience it as two flashes (Shams et al., 2000). Importantly, from our point of view, early cortical processing in one sense modality can be triggered, in the absence of sensory stimulation in this sense modality by cross-modal influences from another sense modality (Calvert et al., 1997; Hertrich et al., 2011; Pekkola et al., 2005; see Nanay, 2018, for further references). These are *bona fide* examples of perceptual processing that are not triggered by corresponding sensory stimulation in this sense modality.

When I am looking at my coffee machine that makes funny noises, I perceive this event by means of both vision and audition. But very often we only receive sensory stimulation from a multisensory event by means of one sense modality. If I hear the noisy coffee machine in the next room, that is, without seeing it, then the question arises: how do I represent the visual aspects of this multisensory event?

The empirical evidence shows that in cases like this one, we have multimodal mental imagery: perceptual processing in one sense modality (here: vision) that is triggered by sensory stimulation in another sense modality (here: audition). Multimodal mental imagery is not a rare and obscure phenomenon. The vast majority of what we perceive are multisensory events: events that can be perceived in more than one sense modality—like the noisy coffee machine. In fact, there are very few events that are not multisensory in this sense. And most of the time, we are only acquainted with these multisensory events via a subset of the sense modalities involved—all the other aspects of these multisensory events are represented by means of multimodal mental imagery.

In discussing music, I will focus on the interplay between the auditory and the visual sense modalities. Listening to music is systematically (and often in an aesthetically relevant manner) influenced by the visual input. A lot of research has focused on how the emotional salience of the visual input influences the perception of music (see Bergeron & Lopes, 2009; Davidson, 1993, for summaries). Instead, I want to talk about how visual input can and often does influence our perception of musical form.

Information in the visual sense modality often highlights or emphasizes the auditory experience of musical form. One simple example of this is the conductor's hand movements that emphasize and highlight certain formal elements of music. Nicolaus Harnoncourt's conducting, with his usually economical movements that only burst into gestures at formally significant points, provides an excellent example. Most of the time, he merely dictates the rhythm—like many other conductors. But occasionally, when something important is happening in the score, he suddenly bursts into an energetic gesture that draws our (visual) attention to what is going on in the musical score at that moment, thereby making the musical form more salient.

Other examples where vision highlights and emphasizes musical form include some ballet and modern dance choreographies, for example, by Mark Morris or Jiri Kylian. Both of these choreographers tend to adjust their choreography to the music in a (sometimes almost comically) synchronous manner. Take Jiri Kylian's choreography *Birthday* for the Nederlands Dans Theater (2006) that uses the music of Mozart's Ouverture of *Le Nozze di Figaro*. Everything the two dancers do in the kitchen (sneeze, cut the dough, break eggs, etc.) is synchronous with the most important musical features—this often leads to comical effects. This choreography makes the musical features that are accompanied by synchronous visual impulses much more salient.

But vision does not always serve to emphasize and highlight musical form. Often, it does the exact opposite: it serves as a counterpoint. Take the famous performance of Rameau's *Les Indes Galantes* by Les Arts Florissants, conducted by William Christie and choreographed by Blanca Li and Andrei Serban (2004, Opera National de Paris). The choreography of the duet *Forêts paisibles* in the last act between Zima and Adario involves very pointed visual gestures against the beat, which makes our multimodal experience of this performance of the duet shift time signature. We hear it as having the time signature of 4/4 instead of the original *alla breve* time signature (2/2) as prescribed in Rameau's score. Here what we see (gestures against the beat) makes us experience the formal properties of the music differently.

To turn to modern dance, some of Pina Bausch's choreographies use the same effect (see Nanay, forthcoming, for a more detailed analysis). At the beginning of her *Café Müller* (1978, Tanztheater Wuppertal), the woman's movements almost always seem to be the exact opposite of what is happening in the musical score (of "O let me weep" from Purcell's *The Fairy Queen*). She stands still for a long time and then suddenly, when there is a lull in the music, starts running; she makes frantic complicated gestures while the music is slower and hardly moves when the music gets faster. The same applies to Bausch's choreography for Gershwin's *The man I love* in her *Nelken* (1982, Tanztheater Wuppertal), where the man's gestures are supposed to express the same meaning as the song's lyrics, but their timing is almost always against the beat. In this interesting example, the auditory experience of both the musical form and the expressive content is influenced by visual effects (see Nanay, 2012, for more examples of this kind).

All of these examples are examples of the multimodal perception of music. But if you experience one of the performances described in the last four paragraphs, your next encounter with this piece of music will be coloured by your multimodal mental imagery. Take the example of Rameau's *Les Indes Galantes* by Les Arts Florissants. If you watch the duet *Forêts paisibles* in the last act between Zima and Adario and then, a couple of days or weeks later you listen to the music alone, because of your prior experience of the visual input against the beat, you will have visual imagery in those moments (influenced by your past experience).

And this multimodal mental imagery has a significant influence on your musical experience (of the music only). Importantly, it will make your experience (of the music only) have the time signature of 4/4 instead of the original *alla breve* time signature (2/2) as prescribed in Rameau's score.

You may say that when you listen to the music and you introspect, you don't feel any multimodal mental imagery. It should be clear from the discussion of mental imagery in general and multimodal mental imagery in particular in the first half of this chapter that mental imagery does not have to be conscious, and it is, in fact, most often unconscious. The same goes for the multimodal mental imagery that plays a role in our perception of music. Even though you are not aware of the multimodal mental imagery when listening to the Rameau duet, your visual cortices are nonetheless active and they also actively influence your auditory cortices in those moments that correspond to the gestures against the beat in the performance.

In fact, multimodal mental imagery here works in a fairly complex manner. The auditory sensory stimulation (of listening to the duet) triggers the visual mental imagery of the gestures against the beat, which in turn influences the auditory processing and musical expectations, which then influences even global characteristics of the musical experience as musical metre. The overall conscious musical experience will be different, but none of these ingredients needs to be conscious. Multimodal mental imagery often works under the radar, but its impact is nonetheless very significant on the overall musical experience.

Part 4: Conclusion

Some purists try to screen out all non-auditory stimuli when listening to music—they close their eyes, for example. This might not be such a great idea, because given the deep multimodality of perception, the multimodal aspects of music perception can enrich our music perception. I gave some examples in the last section of how this could happen (see Nanay, 2015, 2016).

The importance of multimodal mental imagery highlights just how crucial these multimodal aspects are or at least can be. Your multimodal experience of a musical piece even just once can significantly alter not just that one experience but all your experiences of this piece in at least the near future. You are likely to be stuck with the specific multimodal mental imagery after that one instance of listening.

This can be and often is a good thing. But it is also potentially dangerous. Just as listening to one bad performance of a musical piece can have a negative impact on all your subsequent experiences of this piece (because it triggers the wrong kind of musical expectations), listening to a performance of a musical piece with the wrong or distracting multimodal input can also have a negative impact on your subsequent listening, because it will evoke the wrong or distracting multimodal mental imagery.

Multimodal mental imagery evoked in musical listening is aesthetically relevant. But aesthetic relevance can have a positive or a negative valence. Our expectations are malleable, and our multimodal expectations are even more malleable. Again, this can be a good thing or a bad thing. This means that the conclusion of this chapter is not merely of academic importance. It is an important warning to all of us about paying extra attention to how we listen to music.[1]

Note

1 This work was supported by the ERC Consolidator grant (726251), the FWF-FWO grant [G0E0218N], the FWO research grant [G0C7416N], and the FWO-SNF research grant [G025222N]. Special thanks for comments by the two referees, Adriana Renero and Tom Cochrane, and also by the editors of the volume.

References

Atlas, L. Y., & Wager, T. D. (2012). How expectations shape pain. *Neuroscience Letters*, *520*(2), 140–148. https://doi.org/10.1016/j.neulet.2012.03.039

Bergeron, V., & Lopes, D. M. (2009). Hearing and seeing musical expression. *Philosophy and Phenomenological Research*, *78*(1), 1–16. https://doi.org/10.1111/j.1933-1592.2008.00230.x

Bertelson, P. (1999). Ventriloquism: A case of cross-modal perceptual grouping. In G. Aschersleben, T. Bachmann, & J. Müsseler (Eds.), *Cognitive contributions to the perception of spatial and temporal events* (pp. 347–362). Amsterdam: Elsevier.

Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williamns, S. C. R., McGuire, P. K., Woodruff, P. W. R., Iversen, S. D., & David, A. S. (1997). Activation of auditory cortex during silent lipreading. *Science*, *276*(5312), 593–596. https://doi.org/10.1126/science.276.5312.593

Davidson, J. W. (1993). Visual perception of performance manner in the movements of solo musicians. *Psychology of Music*, *21*(2), 103–113. https://doi.org/10.1177/030573569302100201

Ekman, M., Kok, P., & de Lange, F. P. (2017). Time-compressed preplay of anticipated events in human primary visual cortex. *Nature Communications*, *8*(1), 1–9. https://doi.org/10.1038/ncomms15276

Goffaux, P., Redmond, W. J., Rainville, P., & Marchand, S. (2007). Descending analgesia-when the spine echoes what the brain expects. *Pain*, *130*(1–2), 137–143. https://doi.org/10.1016/j.pain.2006.11.011

Herholz, S. C., Halpern, A. R., & Zatorre, R. J. (2012). Neuronal correlates of perception, imagery, and memory for familiar tunes. *Journal of Cognitive Neuroscience*, *24*(6), 1382–1397. https://doi.org/10.1162/jocn_a_00216

Hertrich, I., Dietrich, S., & Ackermann, H. (2011). Cross-modal interactions during perception of audiovisual speech and nonspeech signals: an fMRI study. *Journal of Cognitive Neuroscience*, *23*(1), 221–237. https://doi.org/10.1162/jocn.2010.21421

Jensen, J., McIntosh, A. R., Crawley, A. P., Mikulis, D. J., Remington, G., & Kapur, S. (2003). Direct activation of the ventral striatum in anticipation of aversive stimuli. *Neuron*, *40*(6), 1251–1257. https://doi.org/10.1016/S0896-6273(03)00724-4

Keltner, J. R., Furst, A., Fan, C., Redfern, R., Inglis, B., & Fields, H. L. (2006). Isolating the modulatory effect of expectation on pain transmission: A functional magnetic resonance imaging study. *Journal of Neuroscience*, *26*(16), 4437–4443. https://doi.org/10.1523/jneurosci.4463-05.2006

Kentridge, R. W., Heywood, C. A., & Weiskrantz, L. (1999). Attention without awareness in blindsight. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *266*(1430), 1805–1811. https://doi.org/10.1098/rspb.1999.0850

Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006). *The case for mental imagery*. Oxford: Oxford University Press.

Kouider, S., & Dehaene, S. (2007). Levels of processing during non-conscious perception: A critical review of visual masking. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1481), 857–875. https://doi.org/10.1098/rstb.2007.2093

Koyama, T., McHaffie, J. G., Laurienti, P. J., & Coghill, R. C. (2005). The subjective experience of pain: Where expectations become reality. *Proceedings of the National Academy of Sciences*, *102*(36), 12950–12955. https://doi.org/10.1073/pnas.0408576102

Kraemer, D. J. M., Macrae, C. N., Green, A. E., & Kelley, W. M. (2005). Musical imagery: Sound of silence activates auditory cortex. *Nature*, *434*(7030), 158. https://doi.org/10.1038/434158a

Leaver, A. M., Van Lare, J., Zielinski, B., Halpern, A. R., & Rauschecker, J. P. (2009). Brain activation during anticipation of sound sequences. *The Journal of Neuroscience*, *29*(8): 2477–2485. https://doi.org/10.1523/jneurosci.4921-08.2009

Nanay, B. (2012). The multimodal experience of art. *British Journal of Aesthetics*, *52*(4), 353–363. https://doi.org/10.1093/aesthj/ays042

Nanay, B. (2015). Perceptual content and the content of mental imagery. *Philosophical Studies*, *172*(7), 1723–1736. https://doi.org/10.1007/s11098-014-0392-y

Nanay, B. (2016). *Aesthetics as Philosophy of Perception*. Oxford: Oxford University Press.

Nanay, B. (2017). Pain and mental imagery. *The Monist*, *100*(4), 485–500. https://doi.org/10.1093/monist/onx024

Nanay, B. (2018). Multimodal mental imagery. *Cortex*, *105*, 125–134. https://doi.org/10.1016/j.cortex.2017.07.006

Nanay, B. (2021). Unconscious mental imagery. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1817), 20190689. https://doi.org/10.1098/rstb.2019.0689

Nanay, B. (2023). *Mental imagery*. Oxford: Oxford University Press.

Nanay, B. (forthcoming). Frissons in dance. *Journal of Aesthetics and Art Criticism*.

Pearson, J., Naselaris, T., Holmes, E. A., & Kosslyn, S. M. (2015). Mental imagery: Functional mechanisms and clinical applications. *Trends in Cognitive Sciences*, *19*(10), 590–602. https://doi.org/10.1016/j.tics.2015.08.003

Peerdeman, K. J., van Laarhoven, A. I., Keij, S. M., Vase, L., Rovers, M. M., Peters, M. L., & Evers, A. W. (2016). Relieving patients' pain with expectation interventions: A meta-analysis. *Pain*, *157*(6), 1179–1191. https://doi.org/10.1097/j.pain.0000000000000540

Pekkola, J., Ojanen, V., Autti, T., Jääskeläinen, I. P., Möttönen, R., Tarkiainen, A., & Sams, M. (2005). Primary auditory cortex activation by visual speech: An fMRI study at 3 T. *Neuroreport*, *16*(2), 125–128. https://doi.org/10.1097/00001756-200502080-00010

Ploghaus, A., Tracey, I., Gati, J. S., Clare, S., Menon, R. S., Matthews, P. M., & Rawlins, J. N. P. (1999). Dissociating pain from its anticipation in the human brain. *Science*, *284*(5422), 1979–1981. https://doi.org/10.1126/science.284.5422.1979

Porro, C. A., Baraldi, P., Pagnoni, G., Serafini, M., Facchin, P., Maieron, M., & Nichelli, P. (2002). Does anticipation of pain affect cortical nociceptive systems? *Journal of Neuroscience*, *22*(8), 3206–3214. https://doi.org/10.1523/jneurosci.22-08-03206.2002

Shams, L., Kamitani, Y., & Shimojo, S. (2000). What you see is what you hear. *Nature*, *408*(6814), 788. https://doi.org/10.1038/35048669

Spence, C., & Driver, J. (Eds.). (2004). *Crossmodal space and crossmodal attention*. Oxford: Oxford University Press.

Vroomen, J., Bertelson, P., & de Gelder, B. (2001). Auditory-visual spatial interactions: Automatic versus intentional components. In B. de Gelder, E. de Haan, & C. Heywood (Eds.), *Out of mind* (pp. 140–150). Oxford: Oxford University Press.

Wager, T. D., Rilling, J. K., Smith, E. E., Sokolik, A., Casey, K. L., Davidson, R. J., . . . Cohen, J. D. (2004). Placebo-induced changes in FMRI in the anticipation and experience of pain. *Science*, *303*(5661), 1162–1167. https://doi.org/10.1126/science.1093065

Watkins, S., Shams, L., Tanaka, S., Haynes, J. D., & Rees, G. (2006). Sound alters activity in human V1 in association with illusory visual perception. *Neuroimage*, *31*(3), 1247–1256. https://doi.org/10.1016/j.neuroimage.2006.01.016

Weiskrantz, L. (2009). *Blindsight: A case study spanning 35 years and new developments* (2nd ed.). Oxford: Oxford University Press.

Yokosawa, K., Pamilo, S., Hirvenkari, L., Hari, R., & Pihko, E. (2013). Activation of auditory cortex by anticipating and hearing emotional sounds: An MEG study. *PLoS One*, *8*(11): e80284. https://doi.org/10.1371/journal.pone.0080284

Zatorre, R. J., & Halpern, A. R. (2005). Mental concert: Musical imagery and auditory cortex. *Neuron*, *47*(1), 9–12. https://doi.org/10.1016/j.neuron.2005.06.013

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Part II

Measurement

6 Music-Evoked Imagery and Imagery for Music

Subjective and Behavioural Measures

Rebecca W. Gelding, Robina A. Day, and William Forde Thompson

Music-related visual and auditory imagery are primarily *subjective* experiences that are often richly detailed and idiosyncratic. Therefore, it has been challenging to devise appropriate measures to investigate the nature, characteristics, and cognitive function of these phenomena. This chapter describes some of the subjective and behavioural measures currently being used to investigate two distinct types of imagery that arise in relation to music. The first type refers to imagery that is triggered by a music experience, such as visual imagery for landscapes or other non-musical scenes while listening to music. In this section, we focus on the temporal dynamics of such music-evoked visual imagery and their implications for underlying mechanisms. The second type refers to auditory imagery for sounded music itself, which can arise in the absence of a musical trigger, but may also function to fill in perceptual gaps. Here, we discuss some of the complications of measuring auditory imagery for music, and describe recent advances in this field of study. Other forms of imagery arising from music that are not included in this chapter are kinaesthetic (Godøy, 2001; Kim, this volume), spatial (Eitan & Granot, 2006), and notation-evoked (Wolf et al., 2018).

Goals and Challenges of Measuring Music-Evoked Visual Imagery

A basic goal of measuring music-evoked visual imagery is to estimate the prevalence of its occurrence among music listeners. Another goal is to characterize the nature, content, and quality of visual imagery experiences. What kinds of images do people experience? Are certain features of music more likely to induce imagery? Is imagery static, film-like, vivid, or vague? How long does it take for music to trigger imagery? Such questions require distinct measures, including sensitive measures of the timing, quality, vividness, intensity, and duration of imagery.

Currently, there are no standardized tools to measure music-evoked visual imagery and no systematic classification of the characteristics of listener experiences (Taruffi & Küssner, 2019). One reason for this research gap is that subjective phenomena are difficult to measure. Three common psychological measures are self-report, behavioural, and physiological. Self-report measures consist of

DOI: 10.4324/9780429330070-9

oral or written statements such as open-ended responses, questionnaires, and rating scales (see Hubbard, this volume); behavioural measures include reaction-time and accuracy of response; physiological measures include skin conductance, heart-rate, blood pressure, and measures of electrical activity and blood flow. Recent investigations of music-related imagery have employed a combination of these measures (Day & Thompson, 2019; Taruffi et al., 2017), but many studies are limited to self-reports.

Early investigators collected open-ended qualitative descriptions of imagery in response to musical stimuli (Osborne, 1981; Quittner & Glueckauf, 1983). Such an approach provides rich details, but descriptions are idiosyncratic, and it is difficult to assess their reliability given that participants may elaborate or fill in gaps during the reporting phase. More recent investigations feature questionnaires designed to probe specific aspects of music-evoked visual imagery, such as the prevalence and content of imagery (Küssner & Eerola, 2019), and the impact of imagery on emotional responses (Balteş & Miu, 2014) and aesthetic appreciation (Belfi, 2019). However, biases arising from demand characteristics (the unconscious modification of participant responses to meet the expectations of the research) (Hubbard, 2018), social desirability (trying to impress the experimenter and not reporting honestly), and reactivity (the alteration of participant behaviour due to awareness of being observed) may influence questionnaire outcomes.

Chronometric Measures of Music-Evoked Visual Imagery

Chronometry—the science of the measurement of time—is another form of measurement that has proven both challenging and revealing in recent studies of auditory and visual imagery. Mental chronometry, or the study of processing speeds, is one of the main forms of measurement used to investigate the content, duration, and temporal unfolding of a wide range of cognitive processes (Kranzler, 2012). Measures of reaction time (RT), generally recorded in milliseconds, are used to determine the time it takes for an individual to respond after the presentation of an auditory or visual stimulus (see Jensen, 2006). Chronometric measures have been important in the exploration of mechanisms underlying short-term memory (Sternberg, 1966), mental rotation (Shepard & Metzler, 1971), and implicit association (Schiller et al., 2016).

Imagery is also associated with emotional responses to music (see Taruffi & Küssner, this volume), and chronometric measures can be used to compare the onset of music-evoked visual imagery with the onset of music-evoked emotions (Day & Thompson, 2019). Juslin and Västfjäll (2008) proposed that emotional responses to music can arise from music-evoked visual imagery. In an investigation of this proposal, Day and Thompson (2019) examined whether music listeners typically form visual imagery prior to their emotional response to the music. Such a pattern would be expected if visual imagery were the sole cause of an emotional response to music. Across three experiments, the researchers compared the time taken for listeners to experience an emotional response to music with the time taken to form visual imagery. Measurements were taken for a range of music, including Classical, Popular, Alternative, Heavy Rock, and World Music. In Experiment 1, participants completed a keypress as soon as they: (1) *recognized an emotion* in the music (condition 1); (2) *felt emotion* in response to the music (condition 2); and (3) *experienced a visual image* (condition 3). Mean response times for recognizing an emotion were significantly shorter than the time taken to feel an emotion, and the time taken to feel an emotion was, in turn, significantly shorter than the time taken to form a visual image. Using the same key-press method, Experiments 2 and 3 confirmed that emotional responses to music were generally faster to emerge than experiences of visual imagery. That is, emotional responses to music typically occurred prior to experiences of music-evoked visual imagery. These findings raise intriguing questions about the causal relationship between music-evoked visual imagery and emotional responses to music, and provide useful data that can be incorporated into Juslin and Västfjäll's (2008) model of music and emotion.

Nevertheless, legitimate questions should be asked about the reliability and validity of this chronometric method. Participants were instructed to complete a keypress as soon as they *experienced* an emotion (in one condition) or an image (in a different condition). Presumably, such experiences would need to reach a certain threshold of awareness in order to trigger a behavioural response. It is difficult to equate the threshold of awareness needed for an *emotional* state to trigger a key press with the threshold of awareness needed for *visual imagery* to trigger a key press. Criterion differences could arise, for example, if keypress responses were based on the formation of a propositional representation of each type of experience (Pylyshyn, 1973). It might be easier to represent an emotional state (e.g., sad) than a complex visual image. Another consideration is that the keypress task and the music may have competed for attention, impacting keypress responses in unknown ways. If participants were highly absorbed in the music, they might forget to make a keypress response or their response could be delayed. Competition for attention may have impacted keypress responses for emotional experiences differently than experiences of visual imagery. Finally, the experimental design was open to the same biases as other subjective designs, such as demand characteristics.

Auditory Imagery for Music

A complementary approach to investigating music and mental imagery is to explore the phenomenon of auditory imagery *for* music. Whether anticipating the next song on an album, finding a song "stuck" in your head, or simply bringing to mind a melody—most people report experiencing the phenomenon of hearing music in their minds in the absence of corresponding sensory input (see Floridou, this volume). Historically, research on imagery for music has emphasized the auditory modality (Halpern, 2003). However, imagery for music is often a multimodal experience (Nanay, this volume), whereby both sound and performance movements are imagined (Clark et al., 2012).

Goals and Challenges of Measuring Auditory Imagery for Music

As with music-evoked visual imagery, a fundamental challenge in studying auditory imagery for music is the design of adequate methods for inducing imagery in a way that can be verified (Hubbard, 2013). Asking participants to describe the strategy that they used to complete an imagery task is one useful approach (Gelding et al., 2015), but subjective reports are susceptible to biases and judgement error. Thus, objective, behavioural measures are essential so that investigators can verify that musical imagery has genuinely occurred (Zatorre & Halpern, 2005).

Auditory imagery is often employed in skill development as "mental practice" (Zatorre & Halpern, 2005), but it is challenging to measure such imagery reliably. Clark and Williamon (2012) investigated the auditory imagery abilities of advanced musicians by measuring timing accuracy of live and imagined performances of a two-minute piece prepared to performance standard. In the imagery condition, participants tapped the beat of the music as they used multimodal mental imagery to reconstruct their performance. Tapping and live performances were recorded for comparison of inter-beat-intervals (IBIs), and standardized measures of imagery vividness and usage were also collected. Correlations between IBIs for live and imagined (tapped) performance conditions were significant for 69% of participants, but predictably low (mean $r = .33$) given the narrow range of expressive deviations across a large number of IBIs (range restriction). Moreover, using tapping to measure imagined performances may have confounded the results, given that tapping is itself a form of explicit performance, and dependent on precise motor control to represent the imagined performance.

For many tests of imagery ability, cognitive "short-cuts" can be employed to achieve an accurate response, making it difficult to interpret results. For example, repeatedly imagining just one note (absolute pitch memory) might allow a participant to correctly identify the key signature of a musical stimulus, rather than imagining the full set of notes in a target sequence (Cebrian & Janata, 2010). On the other hand, complex imagery tasks such as mentally reversing whole melodies can be too difficult even for accomplished musicians (Zatorre et al., 2010). Musical training and experience have a profound effect on the ability to perform imagery tasks, and yet measures of musical imagery are not automated to adapt to individual abilities. Adaptable measures will allow investigators to evaluate musical imagery abilities across participants with a range of musical training backgrounds.

Imagery for music involves the generation, maintenance, and manipulation of a musical image, but existing research has primarily focused on the processes of generation and maintenance. There is much less research on the ability to *manipulate* musical imagery, such as imagining a familiar tune played on a xylophone or sung at a higher pitch (but see Greenspon et al., 2017). In short, behavioural paradigms are needed that include objective measures of all three processes of musical imagery, cater for a range of individual abilities, and confirm that an auditory imagery strategy has been used to complete the task.

Auditory Imagery for Songs

The experience of imaging a song occurs in a variety of contexts, and can be classified using various terminology. Effortful or voluntary imagery is the conscious bringing to mind of a piece of music. Many studies of this type of auditory imagery involve familiar songs. For example, participants may be asked to imagine missing parts of songs, and then judge the tuning or timing of the continuation of the song when it returns (Herholz et al., 2008; Weir et al., 2015).

Imagery for music can also occur spontaneously in different contexts. When the context is music listening, the term *anticipatory imagery* is often used, as participants experience the musical image for the continuation of a musical stimulus (Gabriel et al., 2016). Anticipatory imagery may also occur following an incomplete chord progression (Otsuka et al., 2008), or anticipating the next song during the silent gap between songs on a familiar album (Leaver et al., 2009).

Not all spontaneous imagery occurs within a musical listening context, and several types of involuntary musical imagery (INMI) have been described (Floridou, this volume; Hubbard, 2019b). Melodies or fragments of songs that spontaneously and repeatedly occur in the mind are often called "earworms" (Farrugia et al., 2015). Earworms and other forms of INMI often occur after events such as recent music exposure or a memory trigger, or during internal states that are emotional or require low attention (Williamson et al., 2012). Interestingly, the tempo of songs that are spontaneously or voluntarily imagined tends to be relatively stable across different imagery experiences (Jakubowski et al., 2018). However, people who report more INMI experiences are no better at voluntary musical imagery tasks than those who report fewer INMI experiences (Weir et al., 2015).

An important consideration in the examination of auditory imagery for songs is that the content of imagery is variable, and yet studies often require participants to be familiar with specific songs. In addition, songs incorporate several properties of the musical stimuli (e.g., pitch, contour, harmony, rhythm, and lyrics). To understand the independent and joint contributions of musical properties, strategies are needed to isolate these features of the auditory image.

Pitch Imagery

One paradigm that addresses the aforementioned challenges is the Pitch Imagery Arrow Task (PIAT; Gelding et al., 2015). Given a tonal context and an initial pitch sequence, arrows are displayed to elicit a scale-step sequence of imagined pitches, and participants indicate whether the final imagined tone matches an audible probe. Competent task performance requires active musical imagery and is very difficult to achieve using alternative cognitive strategies (Gelding et al., 2015).

The PIAT offers a number of advantages over previous measures of imagery for music. First, the task provides concrete behavioural evidence that participants were using musical imagery. Secondly, each trial is a novel sequence of pitch steps that cannot be anticipated in advance. Thirdly, the PIAT employs a range of difficulty levels that can accommodate individual differences in musical experience and imagery ability: both musicians and non-musicians can complete the task. Fourthly, participants are asked to confirm what strategy they used to complete the task. Finally, the task can be easily adapted to generate comparable control conditions, such as mental arithmetic.

Performance on the PIAT is positively correlated with auditory imagery vividness and mental control, as measured by an established psychometric measure of general auditory imagery ability, the Bucknell Auditory Imagery Scale (BAIS; Halpern, 2015), as well as musical training (Gelding et al., 2015). Additionally, performance on the PIAT has been linked to ability to tap in time to expressive music (Colley et al., 2018) and sing in tune (Greenspon & Pfordresher, 2019). A new adaptive version of the PIAT (aPIAT) efficiently adapts to an individual's imagery performance and hence can be used to test a wider range of imagery ability (Gelding et al., 2021). Performance on the aPIAT showed positive moderate to strong correlations with measures of non-musical and musical working memory, self-reported musical training, and general musical sophistication (see also online demonstrations).

Rhythm Imagery

Whereas the aPIAT involves variable pitch and constant rhythm in a manipulation paradigm, the Rhythm Imagery Task (RIT) employs a constant pitch and variable rhythm (Gelding, 2019). This measure of imagery for rhythm requires participants to mentally maintain simple rhythms that are timed by a bass drum metronome for a continuation period lasting more than 9.6 seconds. Participants then judge the timing of an unpredictable tone as either "in time" or "not in time" with the rhythm. Performance on the RIT was significantly improved after a block of trials in which participants explicitly tapped out the rhythmic patterns, even when auditory feedback from these taps was masked by white noise. This finding suggests that engagement of the motor system enhances imagery performance beyond the benefits of auditory feedback from the tapped rhythm (Highben & Palmer, 2004; Manning et al., 2017; Manning & Schutz, 2015). Tapping accuracy was the strongest predictor of subsequent rhythm imagery performance, again suggesting that the brain's motor system is integral to rhythm imagery. Rhythm imagery accuracy and tapping accuracy, in turn, were positively correlated with musical training and sophistication subscales of the Goldsmiths Musical Sophistication Index (Gold-MSI; Müllensiefen et al., 2014). However, unlike the PIAT, performance on the RIT was not significantly correlated with the Vividness or Control subscales of the BAIS (Gelding, 2019).

Combining Behavioural and Neuroscientific Measures

In order to fully understand the implications of behavioural measures of both music-evoked visual imagery and imagery for music, it may be useful to supplement such measures with neuroscientific measures (see Belfi, this volume). In music-evoked visual imagery, multiple modalities such as questionnaires and behavioural tasks used in conjunction with fMRI have been instrumental in identifying the role of the Default Mode Network (DMN; Taruffi et al., 2017). Specifically within the DMN, the Medial Prefrontal Cortex (MPFC) is strongly activated in Music-Evoked Autobiographical Memories (MEAMs; Janata, 2009), which have a visual imagery component. Future studies should therefore consider combining neuroimaging techniques with temporal measures such as those utilized by Day and Thompson (2019).

The neural mechanisms of imagery for music have been studied since the 1990s, and steady progress has been made using lesion studies (Zatorre & Halpern, 1993) and a variety of neuroimaging techniques (see Hubbard, 2010, 2019a). These studies have found auditory imagery for music involves many, but not all, of the brain areas involved in auditory perception (Albouy et al., 2017; Foster et al., 2013; Kraemer et al., 2005; Müller et al., 2013; Zatorre & Halpern, 2005). For example, recent work has examined network interactions between auditory and sensorimotor regions to understand the nature of coordination between these regions in perception (Ross et al., 2017) and during pitch imagery (Gelding et al., 2019). The strength of such approaches relies on the combination of behavioural and neuroscientific measures to address the questions at hand (Hubbard, 2019a; Zatorre & Halpern, 2005).

Conclusion

In order to progress in our understanding of music-evoked visual imagery and imagery for music, new tasks are needed to assess whether specific music characteristics such as melody, harmony, rhythm, or timbre are associated with the onset of imagery. Well-designed investigations of musical imagery should (1) involve objective measures of performance; (2) be adaptable and flexible to cater for a range of ability levels; (3) confirm the modality of imagery strategy used to complete a task; and (4) use participants with a range of musical abilities to be able to generalize the findings. Finally, the use of multiple tools such as questionnaires, behavioural studies, and neuroimaging studies is recommended in the exploration of both music-evoked visual imagery and auditory imagery for music.

References

Albouy, P., Weiss, A., Baillet, S., & Zatorre, R. J. (2017). Selective entrainment of theta oscillations in the dorsal stream causally enhances auditory working memory performance. *Neuron*, *94*(1), 193–206. doi:10.1016/j.neuron.2017.03.015

Balteş, F. R., & Miu, A. C. (2014). Emotions during live music performance: Links with individual differences in empathy, visual imagery, and mood. *Psychomusicology: Music, Mind, and Brain*, *24*(1), 58–65. doi:10.1037/pmu.0000030

Belfi, A. M. (2019). Emotional valence and vividness of imagery predict aesthetic appeal in music. *Psychomusicology: Music, Mind & Brain*, *29*(2–3), 128–135. doi:10.1037/pmu0000232

Cebrian, A. N., & Janata, P. (2010). Electrophysiological correlates of accurate mental image formation in auditory perception and imagery tasks. *Brain Research*, *1342*, 39–54. doi:10.1016/j.brainres.2010.04.026

Clark, T., & Williamon, A. (2012). Imagining the music: Methods for assessing musical imagery ability. *Psychology of Music*, *40*(4), 471–493. doi:10.1177/0305735611401126

Clark, T., Williamon, A., & Aksentijevic, A. (2012). Musical imagery and imagination: The function, measurement and application of imagery skills for performance. In D. Miell, D. Hargreaves, & R. MacDonald (Eds.), *Musical imaginations: Multidisciplinary perspectives on creativity, performance and perception* (pp. 351–365). Oxford: Oxford University Press.

Colley, I. D., Keller, P. E., & Halpern, A. R. (2018). Working memory and auditory imagery predict sensorimotor synchronization with expressively timed music. *The Quarterly Journal of Experimental Psychology*, *71*(8), 1781–1796. doi:10.1080/17470218.2017.1366531

Day, R. A., & Thompson, W. F. (2019). Measuring the onset of experiences of emotion and imagery in response to music. *Psychomusicology: Music, Mind & Brain*, *29*(2–3), 75–89. doi:10.1037/pmu0000220

Eitan, Z., & Granot, R. Y. (2006). How music moves: Musical parameters and listeners images of motion. *Music Perception: An Interdisciplinary Journal*, *23*(3), 221–248. doi:10.1525/mp.2006.23.3.221

Farrugia, N., Jakubowski, K., Cusack, R., & Stewart, L. (2015). Tunes stuck in your brain: The frequency and affective evaluation of involuntary musical imagery correlate with cortical structure. *Consciousness and Cognition*, *35*, 66–77. doi:10.1016/j.concog.2015.04.020

Foster, N. E. V., Halpern, A. R., & Zatorre, R. J. (2013). Common parietal activation in musical mental transformations across pitch and time. *NeuroImage*, *75*, 27–35. doi:10.1016/j.neuroimage.2013.02.044

Gabriel, D., Wong, T. C., Nicolier, M., Giustiniani, J., Mignot, C., Noiret, N., . . . Vandel, P. (2016). Don't forget the lyrics! Spatiotemporal dynamics of neural mechanisms spontaneously evoked by gaps of silence in familiar and newly learned songs. *Neurobiology of Learning and Memory*, *132*, 18–28. doi:10.1016/j.nlm.2016.04.011

Gelding, R. W. (2019). *Auditory-sensorimotor brain function during mental imagery of musical pitch and rhythm* [Doctoral dissertation, Macquarie University]. Macquarie University Research Online. www.researchonline.mq.edu.au/vital/access/manager/Repository/mq:71110

Gelding, R. W., Harrison, P. M. C., Silas, S., Johnson, B. W., Thompson, W. F., & Müllensiefen, D. (2021). An efficient and adaptive test of auditory mental imagery. *Psychological Research*, *85*(3), 1201–1220. doi:10.1007/s00426-020-01322-3

Gelding, R. W., Thompson, W. F., & Johnson, B. W. (2015). The pitch imagery arrow task: effects of musical training, vividness, and mental control. *PLoS One*, *10*(3), e0121809. doi:10.1371/journal.pone.0121809

Gelding, R. W., Thompson, W. F., & Johnson, B. W. (2019). Musical imagery depends upon coordination of auditory and sensorimotor brain activity. *Scientific Reports*, *9*(1), 1–13. doi:10.1038/s41598-019-53260-9

Godøy, R. I. (2001). Imagined action, excitation, and resonance. In R. I. G. H. Jørgensen (Ed.), *Musical imagery* (pp. 237–250). New York: Taylor & Francis.

Greenspon, E. B., & Pfordresher, P. Q. (2019). Pitch-specific contributions of auditory imagery and auditory memory in vocal pitch imitation. *Attention, Perception, & Psychophysics*, *81*(7), 2473–2481. doi:10.3758/s13414-019-01799-0

Greenspon, E. B., Pfordresher, P. Q., & Halpern, A. R. (2017). Pitch imitation ability in mental transformations of melodies. *Music Perception: An Interdisciplinary Journal*, *34*(5), 585–604. doi:10.1525/mp.2017.34.5.585

Halpern, A. R. (2003). Cerebral substrates of musical imagery. In I. Peretz & R. J. Zatorre (Eds.), *The cognitive neuroscience of music* (pp. 217–230). New York: Oxford University Press.

Halpern, A. R. (2015). Differences in auditory imagery self-reported predict neural and behavioral outcomes. *Psychomusicology: Music, Mind, and Brain*, *25*(1), 37–47. doi:10.1037/pmu0000081

Herholz, S. C., Lappe, C., Knief, A., & Pantev, C. (2008). Neural basis of music imagery and the effect of musical expertise. *European Journal of Neuroscience*, *28*(11), 2352–2360. doi:10.1111/j.1460-9568.2008.06515.x

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.

Hubbard, T. L. (2010). Auditory imagery: Empirical findings. *Psychological Bulletin*, *136*(2), 302–329. doi:10.1037/a0018436

Hubbard, T. L. (2013). Auditory aspects of auditory imagery. In S. Lacey & R. Lawson (Eds.), *Multisensory imagery* (pp. 51–76). New York: Springer.

Hubbard, T. L. (2018). Some methodological and conceptual considerations in studies of auditory imagery. *Auditory Perception & Cognition*, *1*(1–2), 6–41. doi:10.1080/25742442.2018.1499001

Hubbard, T. L. (2019a). Neural mechanisms of musical imagery. In M. H. Thaut & D. A. Hodges (Eds.), *The Oxford handbook of music and the brain*. Oxford: Oxford University Press.

Hubbard, T. L. (2019b). Some anticipatory, kinesthetic, and dynamic aspects of auditory imagery. In M. Grimshaw-Aagaard, M. Walther-Hansen, & M. Knakkergaard (Eds.), *The Oxford handbook of sound and imagination* (Vol. 1). Oxford: Oxford University Press.

Jakubowski, K., Bashir, Z., Farrugia, N., & Stewart, L. (2018). Involuntary and voluntary recall of musical memories: A comparison of temporal accuracy and emotional responses. *Memory & Cognition*, *46*(5), 741–756. doi:10.3758/s13421-018-0792-x

Janata, P. (2009). The neural architecture of music-evoked autobiographical memories. *Cerebral Cortex*, *19*(11), 2579–2594. doi:10.1093/cercor/bhp008

Jensen, A. R. (2006). *Clocking the mind: Mental chronometry and individual differences*. Amsterdam: Elsevier.

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, *31*(5), 559–575. doi:10.1017/S0140525X08005293

Kraemer, D. J. M., Macrae, C. N., Green, A. E., & Kelley, W. M. (2005). Musical imagery: Sound of silence activates auditory cortex. *Nature*, *434*(7030), 158. doi:10.1038/434158a

Kranzler, J. H. (2012). Mental chronometry. In N. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 100–102). Berlin: Springer Science + Business Media.

Küssner, M. B., & Eerola, T. (2019). The contents and functions of vivid and soothing visual imagery during music listening: Findings from a survey study. *Psychomusicology: Music, Mind & Brain*, *29*(2–3), 90–99. doi: 10.1037/pmu0000238

Leaver, A. M., Van Lare, J., Zielinski, B., Halpern, A. R., & Rauschecker, J. P. (2009). Brain activation during anticipation of sound sequences. *The Journal of Neuroscience*, *29*(8), 2477–2485. doi:10.1523/jneurosci.4921-08.2009

Manning, F. C., Harris, J., & Schutz, M. (2017). Temporal prediction abilities are mediated by motor effector and rhythmic expertise. *Experimental Brain Research*, *235*(3), 861–871. doi:10.1007/s00221-016-4845-8

Manning, F. C., & Schutz, M. (2015). Movement enhances perceived timing in the absence of auditory feedback. *Timing & Time Perception*, *3*(1–2), 3–12. doi:10.1163/22134468-03002037

Müllensiefen, D., Gingras, B., Musil, J. J., & Stewart, L. (2014). The musicality of non-musicians: An index for assessing musical sophistication in the general population. *PLoS One*, *9*(2), e89642. doi:10.1371/journal.pone.0089642

Müller, N., Keil, J., Obleser, J., Schulz, H., Grunwald, T., Bernays, R.-L., . . . Weisz, N. (2013). You can't stop the music: Reduced auditory alpha power and coupling between auditory and memory regions facilitate the illusory perception of music during noise. *NeuroImage*, *79*, 383–393. doi:10.1016/j.neuroimage.2013.05.001

Osborne, J. W. (1981). The mapping of thoughts, emotions, sensations, and images as responses to music. *Journal of Mental Imagery*, *5*, 133–136.

Otsuka, A., Tamaki, Y., & Kuriki, S. (2008). Neuromagnetic responses in silence after musical chord sequences. *NeuroReport*, *19*(16), 1637–1641. doi:10.1097/WNR.0b013e32831576fa

Pylyshyn, Z. W (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin*, *80*(1), 1–24. doi:10.1007/978-94-010-1193-8_1

Quittner, A., & Glueckauf, R. (1983). The facilitative effects of music on visual imagery: A multiple measures approach. *Journal of Mental Imagery*, *7*(1), 105–120.

Ross, B., Barat, M., & Fujioka, T. (2017). Sound-making actions lead to immediate plastic changes of neuromagnetic evoked responses and induced beta-band oscillations during perception. *The Journal of Neuroscience*, *37*(24), 5948–5959. doi:10.1523/jneurosci.3613-16.2017

Schiller, B., Gianotti, L. R. R., Baumgartner, T., Nash, K., Koenig, T., & Knoch, D. (2016). Clocking the social mind by identifying mental processes in the IAT with electrical neuroimaging. *Proceedings of the National Academy of Sciences*, *113*(10), 2786–2791. doi:10.1073/pnas.1515828113

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, *171*(3972), 701–703. doi:10.1126/science.171.3972.701

Sternberg, S. (1966). High speed scanning in human memory. *Science*, *153*(3736), 652–654. https://doi.org/10.1126/science.153.3736.652

Taruffi, L., & Küssner, M. B. (2019). A review of music-evoked mental imagery: Conceptual issues, relation to emotion, and functional outcome. *Psychomusicology: Music, Mind & Brain*, *29*(2–3), 62–74. doi:10.1037/pmu0000226

Taruffi, L., Pehrs, C., Skouras, S., & Koelsch, S. (2017). Effects of sad and happy music on mind-wandering and the default mode network. *Scientific Reports*, *7*(1), 14396. doi:10.1038/s41598-017-14849-0

Weir, G., Williamson, V. J., & Müllensiefen, D. (2015). Increased involuntary musical mental activity is not associated with more accurate voluntary musical imagery. *Psychomusicology: Music, Mind & Brain*, *25*(1), 48–57. doi:10.1037/pmu0000076

Williamson, V. J., Jilka, S. R., Fry, J., Finkel, S., Müllensiefen, D., & Stewart, L. (2012). How do "earworms" start? Classifying the everyday circumstances of involuntary musical imagery. *Psychology of Music*, *40*(3), 259–284. doi:10.1177/0305735611418553

Wolf, A., Kopiez, R., & Platz, F. (2018). Thinking in music: An objective measure of notation-evoked sound imagery in musicians. *Psychomusicology: Music, Mind, and Brain*, *28*(4), 209–221. doi:10.1037/pmu0000225

Zatorre, R. J., & Halpern, A. R. (1993). Effect of unilateral temporal-lobe excision on perception and imagery of songs. *Neuropsychologia*, *31*(3), 221–232. doi:10.1016/0028-3932(93)90086-F

Zatorre, R. J., & Halpern, A. R. (2005). Mental concerts: Musical imagery and auditory cortex. *Neuron*, *47*(1), 9–12. doi:10.1016/j.neuron.2005.06.013

Zatorre, R. J., Halpern, A. R., & Bouffard, M. (2010). Mental reversal of imagined melodies: A role for the posterior parietal cortex. *Journal of Cognitive Neuroscience*, *22*(4), 775–789. doi:10.1162/jocn.2009.21239

7 Self-Report Measures in the Study of Musical Imagery

Timothy L. Hubbard

One of the oldest methods in the study of mental imagery involves self-report, in which individuals describe or otherwise respond to questions about their imagery (e.g., Galton, 1880). The majority of studies on mental imagery have focused on visual imagery, but within the past few years, there has been an increasing interest in auditory and especially musical imagery (for review, see Hubbard, 2010, 2019a, 2019b). Although there have been significant methodological advances in the instruments and techniques used in studies of auditory or musical imagery (Hubbard, 2018), self-report measures can still provide important and unique information. Self-report measures regarding imagery can be based on an image generated in response to a question or experimental task or on a recollection or judgement of prior imagery and can involve either verbal or non-verbal responses (see also Gelding et al., this volume). Part 1 considers different types of self-report measures used in studies of musical imagery. Part 2 considers dimensions of musical imagery previously examined with self-report measures and how self-report measures in studies of musical imagery are a necessary complement to other types of measures. Part 3 provides a summary and concludes that self-report measures can provide important information in understanding musical imagery and should not be ignored or dismissed.

Part 1: Types of Self-Report Measures

Self-report measures have typically been in the form of verbal responses in which participants report their experience (e.g., questionnaires and experience sampling) or in the form of non-verbal responses in which patterns of accuracy rates and/or response times in experimental tasks or judgements are examined (e.g., key presses).

Verbal Self-Report Measures

Verbal self-report measures typically involve questionnaires and experience sampling, and aspects of such measures relevant to the study of musical imagery are considered. Advantages and disadvantages of such measures, in general (see also Bartram, 2019; Patten, 2017) and more specific to studies of imagery, are considered.

DOI: 10.4324/9780429330070-10

In Musical Imagery

The majority of questionnaires examining mental imagery focused on visual imagery, and when auditory imagery was examined, the focus was usually on imagery of speech or of sound in general rather than on musical imagery (for review of early studies, see White et al., 1977). A list of questionnaires that specifically focus on musical imagery is presented in Table 7.1.[1] Questionnaires involving imagery have typically been used to assess subjective characteristics or properties of imagery, the level or extent of which could then be related to some other variable (e.g., a questionnaire could ask about the vividness of musical imagery, which could then be related to the level of musical training or experience of that participant). There has been debate regarding whether questionnaires treat imagery as a trait or a state (e.g., D'Angiulli et al., 2013); this issue is not yet resolved (Hubbard, 2018), and both trait (e.g., measures of musical aptitude) and state (e.g., effect of musical imagery on an experimental task) possibilities exist. There are several studies of musical imagery that used experience sampling (in which participants going about their daily lives are queried [e.g., via beeper or text message] at random times to report their current experience; e.g., Bailes, 2006, 2007, 2015), and those tend to focus on individual differences or on the frequency of occurrence of musical imagery. The questions in questionnaires or in experience sampling can be open-ended, but they more commonly present a limited number of options

Table 7.1 Questionnaires that focus on musical imagery.

Questionnaire	*Dimensions or Subscales*	*Authors*
Imaginal Processes Inventory	One factor involves non-voice (music) auditory imagery in daydreaming (see Giambra, 1980)	Singer and Antrobus (1970)
Musical Imagery Questionnaire	Spontaneous musical imagery; factors of unconscious, persistent, entertainment, completeness, musicianship, and distraction	Wammes and Barušs (2009)
Involuntary Musical imagery Scale	Individual differences in involuntary musical imagery; factors of negative valence, movement, personal reflection, and help	Floridou et al. (2015)
Phenomenology of Inner Music questionnaire	Frequency, duration, detail (e.g., melody, harmony, intensity, and presence of instruments and/or lyrics), familiarity, effects on mood and behaviour, perceived control, triggers, mistaken for an external stimulus	Moseley et al. (2018)

(e.g., a set of categorically different alternatives or a Likert-type rating scale) from which participants choose a response.

Advantages

There are several potential advantages to questionnaires. They offer a possibility of (indirect) access to information or processes not directly observable by researchers. Questionnaires are relatively inexpensive to produce and administer, and they can be administered by individuals who are not knowledgeable of experimental hypotheses (e.g., for double-blind research) or without an administrator present (e.g., online). In some settings, anonymity or confidentiality regarding participant identity is possible. Questionnaires can facilitate collection of data from a large number of participants in a relatively brief time. This is especially true if questionnaires are administered online, which can also increase the ethnic, geographic, age, or other diversity in the population being sampled (see Wright, 2005). As noted earlier, in most questionnaires, participants choose one option from a set of pre-specified alternatives or provide a numerical rating for each question; limiting responses in this way allows scoring of questionnaires to be quickly and easily completed and statistical analyses to be applied to the responses. Questionnaires can aid in screening larger populations so that a smaller number of participants can be selected for further testing. Experience sampling offers many of the same advantages as questionnaires. These general advantages also hold if questionnaires and experience sampling are used to study imagery.

Disadvantages

There are several potential disadvantages to questionnaires. Participants might be unwilling or unable to introspect the correct answer, and so a plausible or self-serving answer might be confabulated. The sample could be biased (e.g., questionnaires returned in response to a mass [e]mailing might not be representative of the desired population). If participants misinterpret or misunderstand a question or a rating scale, there might not be an opportunity for correction or clarification, and so responses to that item might not be correct. The limitation of responses to pre-specified options allows the possibility that salient or important information not included in the response options might be missed. Questionnaires might not allow formation of rapport that could lead to disclosure of personal or sensitive information, and gestures and other non-verbal behaviours that might be informative are not captured. It can be difficult to determine participants' truthfulness. There might not be an objectively correct answer to which responses can be compared, and as responses reflect participants' beliefs or opinions, consensus across participants does not necessarily result in accurate conclusions. Relatedly, rating criteria can differ across individuals. Questionnaire construction can be time-consuming because of the need to establish validity and reliability. Experience sampling shares many of these disadvantages, as well as the disadvantage that participants might not be having the desired experience at the time of a query.

In addition to these general disadvantages, there are potential disadvantages specific to imagery research. Episodes of imagery might be relatively transient, and questionnaires and experience sampling might be relatively insensitive to previous transient episodes. There are potential confounds of imagery with non-imagery information (e.g., schema-consistent change and confabulation in memory). Participants might not report on their imagery per se but might report on their interpretation (i.e., verbal measures might encourage modification of imagery content or processing by abstract, semantic, or conceptual [non-imagery] processes) or memory of their imagery. Responses to questionnaires and experience sampling involve introspection, but failures of introspection are well documented (e.g., Lyons, 1986; Nisbett & Wilson, 1977). Perhaps the limitation of self-report measures that has received the most attention within imagery literature is susceptibility to demand characteristics. That is, participants believe that they know the goal or purpose of the experiment, and so they change their responses to support (or not support) that goal or purpose (indeed, demand characteristics are also an important consideration for non-verbal self-report measures of imagery; e.g., Intons-Peterson, 1983; Mitchell & Richman, 1980). Another disadvantage is conceptual confusion resulting from proliferation of terminology across different measures (e.g., Hubbard, 2018, lists 11 different terms used to refer to imagery of speech), and this is especially troubling in light of findings that minor changes in wording, format, or context can result in major changes in responses (cf. Schwartz, 1999).

Non-Verbal Self-Report Measures

Although not often discussed as such, non-verbal responses involving key pressing or other behaviours in studies of imagery are a form of self-report measure, as participants in such studies use a differential response to report on a subjective characteristic or property of their imagery. The use of non-verbal responses increases the variety of potential self-report measures. For example, participants could compare pitches of two imaged lyrics and press one of two keys to indicate which lyric was on the higher pitch (Halpern, 1988a), compare an imaged pitch to a perceived pitch and press one of two keys to indicate whether pitches of the image and the percept were the same or different (Hubbard & Stoeckig, 1988), draw a series of horizontal lines at different heights to reflect imaged musical contour (Weber & Brown, 1986), and adjust a metronome (Halpern, 1988b) or finger tapping rate (Jakubowski et al., 2016) to reflect tempo of an imaged melody. Because studies using non-verbal self-report measures usually present (and/or participants are instructed to image) specific stimuli, there is often an objectively correct answer (e.g., imaged and perceived pitches should or should not match), and this allows the use of statistical techniques to analyse the responses (which often involve accuracy rates and/or response times). Thus, non-verbal self-report measures are typically considered more objective than are verbal self-report measures. Non-verbal measures also have an advantage of not involving language (and abstract, semantic, or conceptual processing invoked by language) or a potential recoding of imagery representations as linguistic representations for verbal report.

Part 2: Dimensions and Complements

Several different dimensions of musical imagery previously studied with self-report measures are noted[2]; additionally, suggestions regarding how self-report measures complement other measures in studies of musical imagery, and the need for convergence among measures, are given.

Dimensions of Musical Imagery

Just as music is considered a multidimensional stimulus, so too can musical imagery be considered multidimensional, and both musical dimensions and non-musical dimensions of musical imagery have been examined. Musical dimensions of musical imagery examined with self-report measures include pitch (Gelding et al., 2015; Hubbard & Stoeckig, 1988), timbre (Crowder, 1989; Halpern et al., 2004), contour (Weber & Brown, 1986), tempo (Halpern, 1988b; Jakubowski et al., 2016), loudness (Bailes et al., 2012), expressive timing (Clark & Williamon, 2012), tonality (Vuvan & Schmuckler, 2011), and lyrical content (Moseley et al., 2018). Findings that imagery of specific pitches (Hubbard & Stoeckig, 1988) or timbres (Crowder, 1989) can prime subsequent perceptual processing of pitch and timbre provided early behavioural evidence that musical imagery involves some of the same mechanisms or processes as does music perception. Musical imagery is predominantly auditory, but self-report measures have revealed auditory, kinaesthetic, and visual information in musical imagery (for review, see Hubbard, 2013b, 2018). Several studies used self-report measures to examine individual differences in musical imagery, with the most common individual difference examined being musical background, training, or sophistication (e.g., Aleman et al., 2000; Bailes, 2007; Bishop et al., 2013; Gelding et al., 2015; Keller & Appel, 2010; Müllensiefen et al., 2014). Non-verbal self-report measures of musical imagery have also been incorporated into tests of musical aptitude (e.g., Gordon, 1965; Seashore et al., 1956).

Non-musical dimensions of musical imagery examined with self-report measures include vividness, clarity, valence, and ability to control (transform) subjective qualities of an auditory image (e.g., Cotter et al., 2019; Halpern, 2015; Willander & Baraldi, 2010; but see Floridou et al., 2015; Wammes & Barušs, 2009, for additional possibilities). A non-musical dimension of musical imagery that has attracted recent attention is whether imagery is voluntary or involuntary (e.g., Floridou et al., 2015; Floridou et al., 2017; Moseley et al., 2018). The dimension of imagery most studied is vividness, and many questionnaires define vividness as the extent to which imagery is similar to perceptual experience. Studies of imagery vividness usually focused on visual imagery (e.g., VVIQ, Marks, 1973; VVIQ-2, Marks, 1995) or collapsed across modalities (e.g., Betts QMI, Sheehan, 1967; but see Andrade et al., 2014), but a few studies on vividness of auditory imagery have been reported (e.g., music students report more vivid auditory imagery than do other students; Campos & Fuentes, 2016). Differences between vividness and clarity (Hubbard & Ruppel, 2021; Willander & Baraldi, 2010), and the usefulness of vividness in consideration of auditory imagery (e.g., Stillman & Kemp, 1993), have been debated. Vividness has been suggested to reflect activation level of working memory (Baddeley & Andrade, 2000) and modality-specific cortices (Belardinelli et al., 2009). However, measures of vividness typically collapse across different potential subprocesses (e.g., image generation, maintenance, inspection, and transformation; Lacey & Lawson, 2013), thus making vividness potentially less useful as a parameter in models of imagery.

A Necessary Complement

Hubbard (2018) suggested that there was not a single type of measure that was optimal for all questions regarding auditory imagery, and a similar notion would hold for musical imagery: Different types of measures might be more or less useful in answering different types of questions (e.g., questionnaires are more useful in addressing subjective aspects of musical imagery; behavioural measures are more useful in addressing how musical imagery functions or interacts with other cognitive processes; and physiological measures are more useful in addressing how musical imagery is instantiated in the brain). Such a suggestion is not unique to imagery, however (e.g., Moskowitz, 1986, reviewed the literature involving self-report, peer report, and behavioural methods in the study of personality and concluded that no single method was always preferred, but that different methods should be applied to different purposes). Given current methodological and technical capabilities, the only way to access subjective content of imagery is by self-report measures. Although self-report measures cannot conclusively demonstrate that imagery was generated or used in any given single experimental task, self-reports of imagery that systematically co-vary across experimental tasks or conditions or with physiological data can provide support for claims that imagery was generated and used by participants and that specific neural structures were related to imagery. Also, self-report measures are necessary for examination of potential emergent properties in imagery (e.g., Finke, 1996) or greater generalizability of principles of imagery at a cognitive level (e.g., Fodor, 1974).

The importance of self-report is also demonstrated in research involving brain imaging (e.g., EEG, PET, and fMRI; for review, see Belfi, this volume; Hubbard, 2019a), which is one of the fastest growing areas of research on musical imagery. In many studies in which brain-imaging data hypothesized to relate to musical imagery are acquired, participants are instructed to generate an image or assumed to use imagery in the experimental task, but no verbal or non-verbal self-report measure was collected to support a claim that an image was actually generated or used in the experimental task (e.g., Meyer et al., 2007; Schurmann et al., 2002; Yoo et al., 2001).[3] This issue is not limited to brain-imaging studies of musical imagery but also occurs in brain-imaging studies of other types of auditory imagery (e.g., Bunzeck et al., 2005; Tian & Poeppel, 2010; Wu et al., 2006), as well as in behavioural studies that posit a role for musical imagery (e.g., Keller & Koch, 2008). In such studies, the possibility that a non-imagery representation or strategy was used by participants cannot be ruled out (Hubbard, 2010; Zatorre & Halpern, 2005), and this has been referred to as *representational ambiguity* (Hubbard, 2018). Such ambiguity has been debated within visual imagery literature (e.g., Kosslyn, 1981; Pylyshyn, 1981) but not within musical (or auditory) imagery literature. Self-report measures are a necessary complement to other measures, not only because self-report measures can reveal emergent or general subjective properties of imagery but also because self-report measures can support (or not) claims that participant responses are based on or involve imagery.

Part 3: Summary and Conclusions

A variety of self-report measures have been used in studies of musical imagery. The most common form of self-report involves verbal responses to questionnaires or experience sampling, which has advantages (e.g., larger numbers of participants, ease of data collection, and relatively low cost) and disadvantages (e.g., responses limited to pre-selected options, limitations of memory, and possibility of demand characteristics). Even so, self-report measures are not limited to verbal responses but can involve non-verbal behavioural responses (e.g., pressing one of two keys to indicate a judgement regarding imaged pitch or timbre, adjusting finger tapping rate or a metronome to reflect tempo of an imaged melody). Self-report measures have examined musical (e.g., pitch, timbre, contour, tempo, loudness, expressive timing, and tonality) and non-musical (e.g., clarity, vividness, valence, control, and voluntariness) dimensions of musical imagery. However, given potential disadvantages of verbal self-report measures, it might be suggested that such measures should not be used and that more objective or observational measures (e.g., brain imaging) be used. Indeed, it might cynically be suggested that self-report measures are just something to use until more rigorous biological, neurological, or chemical techniques are developed. Such a suggestion, though, ignores the possibility of emergent properties and greater generalization of principles at a cognitive level, as well as the importance of using a range of different converging measures and the problem of representational ambiguity.

Although there have been significant methodological advances in the study of musical imagery, self-report measures can still make important contributions to imagery research. Self-report measures currently offer the most straightforward (albeit indirect) way to obtain data on subjective experience of imagery and can complement other measures and provide necessary converging evidence that imagery is actually generated or used in a particular experimental task or by a particular neural structure. Consistent with this, studies using a variety of self-report measures suggest that musical imagery can preserve many of the structural and functional properties of musical stimuli and involve many of the mechanisms used in music perception, cognition, and performance (e.g., Hubbard, 2010, 2013a, 2013b, 2019a, 2019b). Additionally, areas of research that are not focused on imagery might also be informed by self-report measures regarding imagery (e.g., musical practice and performance). As long as the subjective experience of imagery cannot be objectively observed or experienced by another individual, self-report measures will remain an important tool in the study of mental imagery. Complementary and converging measures in neurophenomenology (Gallagher, 2009; Varela, 1996) or heterophenomenology (Dennett, 2007), in which phenomenology is combined with neuroscience and other objective or observational methods, might be a promising direction for future research in imagery. Self-report measures would be a necessary and critical complement in such approaches.

Notes

1 There are many other questionnaires that examine auditory imagery but do not focus on musical imagery, and these include the Betts QMI (Sheehan, 1967), Survey of Mental Imagery (Switras, 1978), Auditory Imagery Scale (Campos, 2017; Gissurarson, 1992), Scale for Inner Speech (Siegrist, 1995), Self-Talk Scale (Brinthaupt et al., 2009), Auditory Imagery Questionnaire (Campos, 2017; Hishitani, 2009), Clarity of Auditory Imagery Scale (Willander & Baraldi, 2010), Varieties of Inner Speech Questionnaire (McCarthy-Jones & Fernyhough, 2011), Plymouth Sensory Imagery Questionnaire (Andrade et al., 2014), and Bucknell Auditory Imagery Scale (Halpern, 2015).

2 Contributions of self-report measures of musical imagery in studies of musical practice and performance (Hubbard, 2018; Presicce, this volume), earworms (Beaman, 2018; Floridou, this volume), and neural processing (Belfi, this volume; Hubbard, 2019a) have been recently reviewed elsewhere, and so are not discussed in this chapter.

3 In notable exceptions, activation in auditory association areas during an unexpected silent gap in familiar music was accompanied by self-reports that the melody was experienced as continuing in imagery during the gap (Gabriel et al., 2016; Kraemer et al., 2005), and activity in rostral prefrontal cortex and motor areas during the silent period before an upcoming track on a familiar CD was accompanied by self-reports of imagery of the upcoming track during the silent period (Leaver et al., 2009). Also, participants who self-reported more vivid musical imagery exhibited greater activation in right superior temporal gyrus and prefrontal cortex (Herholz et al., 2012) and right parietal cortex (Zatorre et al., 2010) and exhibited higher grey matter volume in left inferior parietal lobe, medial superior frontal gyrus, middle frontal gyrus, and left supplementary motor area (Lima et al., 2015).

References

Aleman, A., Nieuwenstein, M. R., Böcker, K. B. E., & de Haan, E. H. F. (2000). Music training and mental imagery ability. *Neuropsychologia*, *38*(12), 1664–1668. https://doi.org/10.1016/s0028-3932(00)00079-8

Andrade, J., May, J., Deeprose, C., Baugh, S.-J., & Ganis, G. (2014). Assessing vividness of mental imagery: The Plymouth Sensory Imagery Questionnaire. *British Journal of Psychology*, *105*(4), 547–563. https://doi.org/10.1111/bjop.12050

Baddeley, A. D., & Andrade, J. (2000). Working memory and the vividness of imagery. *Journal of Experimental Psychology: General*, *129*(1), 126–145. https://doi.org/10.1037/0096-3445.129.1.126

Bailes, F. (2006). The use of experience-sampling methods to monitor musical imagery in everyday life. *Musicae Scientiae*, *10*(2), 173–190. https://doi.org/10.1177/102986490601000202

Bailes, F. (2007). The prevalence and nature of imaged music in the everyday lives of music students. *Psychology of Music*, *35*(4), 555–570. https://doi.org/10.1177/0305735607077834

Bailes, F. (2015). Music in mind? An experience sampling study of what and when, towards an understanding of why. *Psychomusicology: Music, Mind, and Brain*, *25*(1), 58–68. https://doi.org/10.1037/pmu0000078

Bailes, F., Bishop, L., Stevens, C. J., & Dean, R. T. (2012). Mental imagery for musical changes in loudness. *Frontiers in Psychology*, *3*, 525. https://doi.org/10.3389/fpsyg.2012.00525

Bartram, B. (2019). Using questionnaires. In M. Lambert (Ed.), *Practical research methods in education: An early researcher's critical guide* (pp. 1–11). Abingdon-on-Thames: Routledge. https://doi.org/10.4324/9781351188395-1

Beaman, C. P. (2018). The literary and recent scientific history of the earworm: A review and theoretical framework. *Auditory Perception & Cognition*, *1*(1–2), 42–65. https://doi.org/10.1080/25742442.2018.1533735

Belardinelli, M. O., Palmiero, M., Sestieri, C., Nardo, D., Di Matteo, R., Londei, A., . . . Romani, G. L. (2009). An fMRI investigation on image generation in different sensory modalities: The influence of vividness. *Acta Psychologica*, *132*(2), 190–200. https://doi.org/10.1016/j.actpsy.2009.06.009

Bishop, L., Bailes, F., & Dean, R. T. (2013). Musical expertise and the ability to imagine loudness. *PLoS One*, *8*(2): e56052. doi:10.1371/journal.pone.0056052

Brinthaupt, T. M., Hein, M. B., & Kramer, T. E. (2009). The self-talk scale: Development, factor analysis, and validation. *Journal of Personality Assessment*, *91*(1), 82–92. https://doi.org/10.1080/00223890802484498

Bunzeck, N., Wuestenberg, T., Lutz, K., Heinze, H. J., & Jancke, L. (2005). Scanning silence: mental imagery of complex sounds. *Neuroimage*, *26*(4), 1119–1127. https://doi.org/10.1016/j.neuroimage.2005.03.013

Campos, A. (2017). A research note on the factor structure, reliability, and validity of the Spanish version of two auditory imagery measures. *Imagination, Cognition, and Personality*, *36*(3), 301–311. https://doi.org/10.1177/0276236616670892

Campos, A., & Fuentes, L. (2016). Musical studies and the vividness and clarity of auditory imagery. *Imagination, Cognition, and Personality*, *36*(1), 75–84. https://doi.org/10.1177/0276236616635985

Clark, T., & Williamon, A. (2012). Imagining the music: Methods for assessing musical imagery ability. *Psychology of Music*, *40*(4), 471–493. https://doi.org/10.1177/0305735611401126

Cotter, K. N., Christensen, A. P., & Silvia, P. J. (2019). Understanding inner music: A dimensional approach to musical imagery. *Psychology of Aesthetics, Creativity, and the Arts*, *13*(4), 489–503. https://doi.org/10.1037/aca0000195

Crowder, R. G. (1989). Imagery for musical timbre. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(3), 472–478. https://doi.org/10.1037/0096-1523.15.3.472

D'Angiulli, A., Runge, M., Faulkner, A., Zakizadeh, J., Chan, A., & Morcos, S. (2013). Vividness of visual imagery and incidental recall of verbal cues, when phenomenological availability reflects long-term memory accessibility. *Frontiers in Psychology*, *4*, 1. https://doi.org/10.3389/fpsyg.2013.00001

Dennett, D. C. (2007). Heterophenomenology reconsidered. *Phenomenology and the Cognitive Sciences*, *6*(1–2), 247–270. https://doi.org/10.1007/s11097-006-9044-9

Finke, R. A. (1996). Imagery, creativity, and emergent structure. *Consciousness and Cognition*, *5*(3), 381–393. https://doi.org/10.1006/ccog.1996.0024

Floridou, G. A., Williamson, V. J., & Stewart, L. (2017). A novel indirect method for capturing involuntary musical imagery under varying cognitive load. *Quarterly Journal of*

Experimental Psychology, *70*(11), 2189–2199. https://doi.org/10.1080/17470218.2016.1227860

Floridou, G. A., Williamson, V. J., Stewart, L., & Müllensiefen, D. (2015). The Involuntary Musical Imagery Scale (IMIS). *Psychomusicology: Music, Mind, and Brain*, *25*(1), 28–36. https://doi.org/10.1037/pmu0000067

Fodor, J. (1974). Special sciences (or: The disunity of science as a working hypothesis). *Synthese*, *28*(2), 97–115. https://doi.org/10.1007/bf00485230

Gabriel, D., Wong, T. C., Nicolier, M., Giustiniani, J., Mignot, C., Noiret, N., . . . Vandel, P. (2016). Don't forget the lyrics! Spatiotemporal dynamics of neural mechanisms spontaneously evoked by gaps of silence in familiar and newly learned songs. *Neurobiology of Learning and Memory*, *132*, 18–28. https://doi.org/10.1016/j.nlm.2016.04.011

Gallagher, S. (2009). Neurophenomenology. In T. Bayne, A. Cleeremans, & P. Wilken (Eds.), *Oxford companion to consciousness* (pp. 470–472). Oxford: Oxford University Press.

Galton, F. (1880). Statistics of mental imagery. *Mind: A Quarterly Review of Psychology and Philosophy*, *5*(19), 301–318. https://doi.org/10.1093/mind/os-v.19.301

Gelding, R. W., Thompson, W. F., & Johnson, B. W. (2015). The pitch imagery arrow task: Effects of musical training, vividness, and mental control. *PLoS One*, *10*(3), e0121809. doi:10.1371/journal. pone.0121809

Giambra, L. M. (1980). A factor analysis of the items of the Imaginal Processes Inventory. *Journal of Clinical Psychology*, *36*(2), 383–409. https://doi.org/10.1002/jclp.6120360203

Gissurarson, L. R. (1992). Reported auditory imagery and its relationship with visual imagery. *Journal of Mental Imagery*, *16*, 117–122.

Gordon, E. (1965). The musical aptitude profile: A new and unique social aptitude test battery. *Bulletin of the Council for Research in Music Education*, *6*, 12–16.

Halpern, A. R. (1988a). Mental scanning in auditory imagery for songs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(3), 434–443. https://doi.org/10.1037/0278-7393.14.3.434

Halpern, A. R. (1988b). Perceived and imagined tempos of familiar songs. *Music Perception*, *6*(2), 193–202. https://doi.org/10.2307/40285425

Halpern, A. R. (2015). Differences in auditory imagery self-report predict neural and behavioral outcomes. *Psychomusicology: Music, Mind, and Brain*, *25*(1), 37–47. https://doi.org/10.1037/pmu0000081

Halpern, A. R., Zatorre, R. J., Bouffard, M., & Johnson, J. A. (2004). Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia*, *42*(9), 1281–1292. https://doi.org/10.1016/j.neuropsychologia.2003.12.017

Herholz, S. C., Halpern, A. R., & Zatorre, R. J. (2012). Neuronal correlates of perception, imagery, and memory for familiar tunes. *Journal of Cognitive Neuroscience*, *24*, 1382–1397. https://doi.org/10.1162/jocn_a_00216

Hishitani, S. (2009). Auditory Imagery Questionnaire: Its factorial structure, reliability, and validity. *Journal of Mental Imagery*, *33*, 63–80.

Hubbard, T. L. (2010). Auditory imagery: Empirical findings. *Psychological Bulletin*, *136*(2), 302–329. https://doi.org/10.1037/a0018436

Hubbard, T. L. (2013a). Auditory aspects of auditory imagery. In S. Lacey & R. Lawson (Eds.), *Multisensory imagery* (pp. 51–76). New York: Springer.

Hubbard, T. L. (2013b). Auditory imagery contains more than audition. In S. Lacey & R. Lawson (Eds.), *Multisensory imagery* (pp. 221–247). New York: Springer.

Hubbard, T. L. (2018). Some methodological and conceptual considerations in studies of auditory imagery. *Auditory Perception & Cognition*, *1*(1–2), 6–41. https://doi.org/10.1080/25742442.2018.1499001

Hubbard, T. L. (2019a). Neural mechanisms of musical imagery. In M. Thaut & D. A. Hodges (Eds.), *The Oxford handbook of music and the brain* (pp. 521–545). New York: Oxford University Press.

Hubbard, T. L. (2019b). Some anticipatory, kinesthetic, and dynamic aspects of auditory imagery. In M. Grimshaw, M. Walther-Hansen, & M. Knakkergaard (Eds.), *The Oxford handbook of sound and imagination* (Vol. 1, pp. 149–173). New York: Oxford University Press.

Hubbard, T. L., & Ruppel, S. E. (2021). Vividness, clarity, and control in auditory imagery. *Imagination, Cognition, and Personality, 41*(2), 207–221. https://doi.org/10.1177/02762366211013468

Hubbard, T. L., & Stoeckig, K. (1988). Musical imagery: Generation of tones and chords. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(4), 656–667. https://doi.org/10.1037/0278-7393.14.4.656

Intons-Peterson, M. J. (1983). Imagery paradigms: How vulnerable are they to experimenters' expectations? *Journal of Experimental Psychology: Human Perception and Performance*, *9*(3), 394–412. https://doi.org/10.1037/0096-1523.9.3.394

Jakubowski, K., Farrugia, N., & Stewart, L. (2016). Probing imagined tempo for music: Effects of motor engagement and musical experience. *Psychology of Music*, *44*(6), 1274–1288. https://doi.org/10.1177/0305735615625791

Keller, P. E., & Appel, M. (2010). Individual differences, auditory imagery, and the coordination of body movements and sounds in musical ensembles. *Music Perception*, *28*(1), 27–46. https://doi.org/10.1525/mp.2010.28.1.27

Keller, P. E., & Koch, I. (2008). Action planning in sequential skills: Relations to music performance. *Quarterly Journal of Experimental Psychology*, *61*(2), 275–291. https://doi.org/10.1080/17470210601160864

Kosslyn, S. M. (1981). The medium and the message in mental imagery: A theory. *Psychological Review*, *88*(1), 46–66. https://doi.org/10.1037/0033-295x.88.1.46

Kraemer, D. J. M., Macrae, C. N., Green, A. E., & Kelley, W. M. (2005). Musical imagery: Sound of silence activates auditory cortex. *Nature*, *434*(7030), 158. https://doi.org/10.1038/434158a

Lacey, S., & Lawson, R. (2013). Imagery questionnaires: Vividness and beyond. In S. Lacey & R. Lawson (Eds.), *Multisensory imagery* (pp. 271–282). New York: Springer.

Leaver, A. M., Van Lare, J., Zielinski, B., Halpern, A. R., & Rauschecker, J. P. (2009). Brain activation during anticipation of sound sequences. *Journal of Neuroscience*, *29*(8), 2477–2485. https://doi.org/10.1523/jneurosci.4921-08.2009

Lima, C. F., Lavan, N., Evans, S., Agnew, Z., Halpern, A. R., Shanmugalingam, P., . . . Scott, S. K. (2015). Feel the noise: Relating individual differences in auditory imagery to the structure and function of sensorimotor systems. *Cerebral Cortex*, *25*(11), 4638–4650. https://doi.org/10.1093/cercor/bhv134

Lyons, W. (1986). *The disappearance of introspection*. Cambridge, MA: MIT Press.

Marks, D. F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology*, *64*(1), 17–24. https://doi.org/10.1111/j.2044-8295.1973.tb01322.x

Marks, D. F. (1995). New directions for mental imagery research. *Journal of Mental Imagery*, *19*, 153–167.

McCarthy-Jones, S., & Fernyhough, C. (2011). The varieties of inner speech: Links between quality of inner speech and psychopathological variables in a sample of young adults. *Consciousness and Cognition*, *20*(4), 1586–1593. https://doi.org/10.1016/j.concog.2011.08.005

Meyer, M., Elmer, S., Baumann, S., & Jancke, L. (2007). Short-term plasticity in the auditory system: Differential neural responses to perception and imagery of speech and music. *Restorative Neurology and Neuroscience*, *25*(3–4), 411–431.

Mitchell, D. B., & Richman, C. L. (1980). Confirmed reservations: Mental travel. *Journal of Experimental Psychology: Human Perception and Performance*, *6*(1), 58–66. https://doi.org/10.1037/0096-1523.6.1.58

Moseley, P., Alderson-Day, B., Kumar, S., & Fernyhough, C. (2018). Musical hallucinations, musical imagery, and earworms: A new phenomenological survey. *Consciousness and Cognition*, *65*, 83–94. https://doi.org/10.1016/j.concog.2018.07.009

Moskowitz, D. S. (1986). Comparison of self-reports, reports by knowledgeable informants, and behavioral observation data. *Journal of Personality*, *54*(1), 294–317. https://doi.org/10.1111/j.1467-6494.1986.tb00396.x

Müllensiefen, D., Fry, J., Jones, R., Jilka, S., Stewart, L., & Williamson, V. (2014). Individual differences predict patterns in spontaneous involuntary musical imagery. *Music Perception*, *31*(4), 323–338. https://doi.org/10.1525/mp.2014.31.4.323

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*(3), 231–259. https://doi.org/10.1037/0033-295x.84.3.231

Patten, M. L. (2017). *Questionnaire research: A practical guide* (4th ed.). New York: Routledge.

Pylyshyn, Z. W. (1981). The imagery debate: Analogue media versus tacit knowledge. *Psychological Review*, *88*(1), 16–45. https://doi.org/10.1037/0033-295x.88.1.16

Schurmann, M., Raji, T., Fujiki, N., & Hari, R. (2002). Mind's ear in a musician: Where and when in the brain. *Neuroimage*, *16*(2), 434–440. https://doi.org/10.1006/nimg.2002.1098

Schwartz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist*, *54*(2), 93–105. https://doi.org/10.1037/0003-066x.54.2.93

Seashore, C. E., Lewis, D., & Saetveit. J. G. (1956). *Seashore measures of musical talents*. New York: Psychological Corporation.

Sheehan, P. W. (1967). A shortened form of Betts' questionnaire upon mental imagery. *Journal of Clinical Psychology*, *23*(3), 386–389. https://doi.org/10.1002/1097-4679(196707)23:3%3C386::aid-jclp2270230328%3E3.0.co;2-s

Siegrist, M. (1995). Inner speech as a cognitive process mediating self-consciousness and inhibiting self-deception. *Psychological Reports*, *76*(1), 259–265. https://doi.org/10.2466/pr0.1995.76.1.259

Singer, J. L., & Antrobus, J. S. (1970). *Imaginal processes inventory*. New York: City University of New York.

Stillman, J., & Kemp, T. (1993). Visual versus auditory imagination: Image qualities, perceptual qualities, and memory. *Journal of Mental Imagery*, *17*, 181–194.

Switras, J. E. (1978). An alternate-form instrument to assess vividness and controllability of mental imagery in seven modalities. *Perceptual & Motor Skills*, *46*(2), 379–384. https://doi.org/10.2466/pms.1978.46.2.379

Tian, X., & Poeppel, D. (2010). Mental imagery of speech and movement implicates the dynamics of internal forward models. *Frontiers in Psychology*, *1*, 166. https://doi.org/10.3389/fpsyg.2010.00166

Varela, F. J. (1996). Neurophenomenology: A methodological remedy for the hard problem. *Journal of Consciousness Studies*, *3*(4), 330–349.

Vuvan, D. T., & Schmuckler, M. A. (2011). Tonal hierarchy representations in auditory imagery. *Memory & Cognition*, *39*(3), 477–490. https://doi.org/10.3758/s13421-010-0032-5

Wammes, M., & Barušs, I. (2009). Characteristics of spontaneous musical imagery. *Journal of Consciousness Studies*, *16*(1), 37–61.

Weber, R. J., & Brown, S. (1986). Musical imagery. *Music Perception*, *3*(4), 411–426. https://doi.org/10.2307/40285346

White, K., Sheehan, P. W., & Ashton, R. (1977). Imagery assessment: A survey of self-report measures. *Journal of Mental Imagery*, *1*(1), 145–169.

Willander, J., & Baraldi, S. (2010). Development of a new Clarity of Auditory Imagery Scale. *Behavior Research Methods*, *42*(3), 785–790. https://doi.org/10.3758/brm.42.3.785

Wright, K. B. (2005). Researching internet-based populations: Advantages and disadvantages of online survey research, online questionnaire authoring software packages, and web survey services. *Journal of Computer-Mediated Communication*, *10*(3), JCMC1034. https://doi.org/10.1111/j.1083-6101.2005.tb00259.x

Wu, J., Mai, X., Chan, C. C. H., Zheng, Y., & Luo, Y. (2006). Event related potentials during mental imagery of animal sounds. *Psychophysiology*, *43*(6), 592–597. https://doi.org/10.1111/j.1469-8986.2006.00464.x

Yoo, S. S., Lee, C. U., & Choi, B. G. (2001). Human brain mapping of auditory imagery: Event-related functional MRI study. *Neuroreport*, *12*(14), 3045–3049. https://doi.org/10.1097/00001756-200110080-00013

Zatorre, R. J., & Halpern, A. R. (2005). Mental concerts: Musical imagery and the auditory cortex. *Neuron*, *47*(1), 9–12. https://doi.org/10.1016/j.neuron.2005.06.013

Zatorre, R. J., Halpern, A. R., & Bouffard, M. (2010). Mental reversal of imagined melodies: A role for the posterior parietal cortex. *Journal of Cognitive Neuroscience*, *22*(4), 775–789. https://doi.org/10.1162/jocn.2009.21239

8 Neuroscience Measures of Music and Mental Imagery

Amy M. Belfi

Methods from cognitive neuroscience have been frequently used to study mental imagery associated with music listening, imagination, and performance. But if the experience of mental imagery is a subjective one, what value is added by adopting neuroscientific methods? Behavioural or self-report measures may be sufficient for understanding when, how, and why a listener experiences vivid imagery evoked by music (see Gelding et al., this volume; Hubbard, this volume). However, neuroscience techniques can complement behavioural and self-report approaches by identifying neural systems associated with these subjective experiences. If the goal of a research programme is understanding how the experience of music-evoked mental imagery is represented in the brain, a combination of behavioural and neuroscience methods is most appropriate.

For the purposes on this chapter, *mental imagery* will be defined as the process of imagining a sensory object without direct sensory stimulation. In many cognitive neuroscience studies of mental imagery, such imagination is often accompanied by activity in sensory-specific cortices. For example, prior neuroimaging work indicates that mental imagery alone can activate visual sensory cortices, suggesting that imagery functions as a "weak form of perception" (Pearson et al., 2015). While neuroscience methods are commonly used to investigate mental imagery and music, there are many factors that influence the choice of measurement technique. This chapter aims to serve as a practical guide for researchers wishing to start using neuroscientific methods in their studies of music and imagery. Here I review the most commonly used neuroscientific methods in work on music and mental imagery: neuropsychology, functional magnetic resonance imaging (fMRI), scalp and intracranial electroencephalography (EEG), and magnetoencephalography (MEG). For each method, a description of the procedures, advantages, and limitations is provided, as well as specific examples of recent studies using the method to investigate music and mental imagery.

Neuropsychology

The neuropsychological approach, or the study of brain–behaviour relationships in patients with neurological damage, long predates neuroimaging and other technological advances in the field of neuroscience. Originating from the field of

DOI: 10.4324/9780429330070-11

neurology, the study of behavioural changes in patients with brain damage (due to stroke, neurosurgery, trauma, or other neurological disorders) has had a large influence over the treatment and rehabilitation of such patients. However, neuropsychological approaches have also provided strong basic science evidence of brain–behaviour relationships. For example, several studies of patients with focal brain lesions have informed knowledge of the neural basis of music perception and emotion recognition (Dellacherie et al., 2011; Gosselin et al., 2011; Samson et al., 2002).

The lesion method is a term used to describe the systematic study of brain–behaviour relationships in patient populations with *focal* brain damage. Typical causes of such focal damage include cerebrovascular disease, such as stroke; resection of brain tissue for benign tumours or medically intractable epilepsy; or infectious diseases, such as herpes simplex encephalitis. Assuming that the researcher is wishing to conduct a hypothesis-driven study, the first step would be identifying a hypothesis for a particular brain–behaviour relationship. For example, a researcher may predict that the medial prefrontal cortex is a critical brain region underlying music-evoked autobiographical memories (Belfi et al., 2018). Next, the researcher must identify a group of patients with relatively selective and focal damage to their region of interest. In addition to the group of interest (called the *target* group), researchers must also identify comparison groups. Comparison groups should be matched to the target group on demographic variables, such as age, sex, and education levels. There are ideally two comparison groups: a *brain-damaged comparison* group, consisting of individuals with focal damage to regions *outside* the target area. This group provides a control for the effects of brain damage more generally. The second comparison group is a *healthy comparison* group, consisting of individuals with no history of neurological damage, to control for the general effects of age, sex, and other demographic variables.

One key advantage of this approach is that it provides a stronger inference for an underlying brain–behaviour relationship: Identifying behavioural or cognitive impairments in a patient with damage to a particular brain region illustrates the *necessity* of that brain region for that behaviour. For example, patients with left temporal polar damage are impaired at naming famous musical melodies. This indicates that the left temporal pole is *necessary* for naming melodies (Belfi & Tranel, 2014). In contrast, neuroimaging techniques alone can only provide correlational evidence for the *involvement* of a brain region in a certain behaviour.

While there are advantages to studying brain–behaviour relationships using a neuropsychological approach, it is not without limitations. One potential drawback is the limited access to patient populations. First, such patients are relatively rare, in comparison to studying healthy adults. The scarcity of participants also depends on which region of the brain is of interest—that is, focal damage to certain areas of the brain is less likely than others (due to the nature of vascular supply to the brain, common locations of brain tumours, etc.). Secondly, unlike lesion work in animal models, in which researchers can induce focal lesions to specific brain regions, human lesion studies must take advantage of naturally occurring brain lesions. Such naturally occurring lesions are often extensive and cover large amounts of brain tissue, making inferences about specific, focal regions, more challenging. A third, and related, limitation of this type of work is that identifying individuals with brain damage likely requires collaboration with neurologists and/or neuropsychologists, and therefore, it is labour- and expertise-intensive work.

Examples of Neuropsychological Studies of Music and Mental Imagery

One example of music-related mental imagery is the phenomenon of musical hallucinations, which are one form of involuntary musical imagery, or the involuntary perception of music when no music is present (Moseley et al., 2018). In contrast to other forms of involuntary musical imagery, such as earworms, which are perceived as coming from "in one's head," musical hallucinations are perceived as being under less voluntary control and originating from the external world (Moseley et al., 2018). One study attempted to identify brain regions that, when damaged, were associated with musical hallucinations in a group of 37 patients with focal brain lesions (Golden & Josephs, 2015). They identified that musical hallucinations were most commonly associated with lesions to the temporal lobes, primarily in the left hemisphere. Similarly, a recent systematic review of neuropsychological studies of musical hallucinations also identified the role of the temporal lobes (Bernardini et al., 2017).

This exploratory work in neuropsychological populations suggests that damage to the temporal lobes is associated with the presence of musical hallucinations. At face value, it may seem paradoxical that damage to temporal lobe regions involved in auditory perception would be associated with an *increase* in auditory symptoms. It might be equally plausible that damage to the temporal lobes would be associated with less, rather than more, auditory imagery. Though the directionality of this effect remains unclear (i.e., whether temporal lobe damage leads to increased or decreased auditory imagery), it is consistent with neuroimaging evidence (described later), suggesting that mental imagery in response to music involves sensory-specific cortices. One suggested mechanism is that damage to structures in the auditory system may lower the threshold for spontaneous activity elsewhere in this network, resulting in auditory hallucinations. Additionally, prior work suggests that nearby structures surrounding the lesioned area may display "compensatory overactivation" which could then lead to auditory hallucinations (Calabrò et al., 2012).

Functional Magnetic Resonance Imaging (fMRI)

Functional magnetic resonance imaging (fMRI) is the most frequently used neuroscience method when investigating mental imagery and music. Developed in the late 20th century, fMRI quickly became one of the most popular neuroimaging methods within cognitive neuroscience. Researchers use fMRI to identify neural regions associated with a wide spectrum of human behaviours, including the relationships between music and mental imagery.

A number of excellent textbooks provide detailed information on the physics of MRI and fMRI data analysis techniques (e.g., Huettel et al., 2014; Poldrack, 2011a). Briefly, MRI scanners use a magnetic field to generate images of the brain. To use an analogy, a standard (or structural) MRI provides the equivalent of a still "picture" of the brain. Functional MRI (fMRI) provides the equivalent of a "movie," depicting activity changes in the brain over time. As opposed to directly measuring neural activity, fMRI measures changes in blood oxygenation (the blood-oxygen-level-dependent or BOLD response). Designing an fMRI task is similar to a behavioural experiment, where participants complete multiple repeated trials of the behaviour of interest while in the MRI scanner. Researchers must design a study that can be presented within the confines of the scanner and recruit subjects who are able to lie still in a confined space for the duration of the experiment.

When discussing advantages and limitations of neuroimaging techniques, it is important to mention the concepts of spatial and temporal resolution. Spatial resolution is the degree to which a measurement technique can be used to *localize* brain activity (i.e., *where* does the activity occur). Temporal resolution is the degree to which the technique can be used to identify the precise *timing* of the activity (i.e., *when* does the activity occur). fMRI has relatively good spatial resolution, at the level of a voxel (3D pixel). Voxel size is determined by the researcher but is typically on the order of cubic millimetres. The temporal resolution of fMRI is relatively slow, as the BOLD response is dependent on blood flow (on the order of seconds). Figure 8.1 depicts various neuroimaging techniques in terms of their spatial and temporal resolution. Researchers should consider their question of interest when choosing the most appropriate neuroimaging technique. For example, if a researcher has a prediction about a specific neural region that is associated with a phenomenon of interest, they should adopt an imaging technique with high spatial resolution.

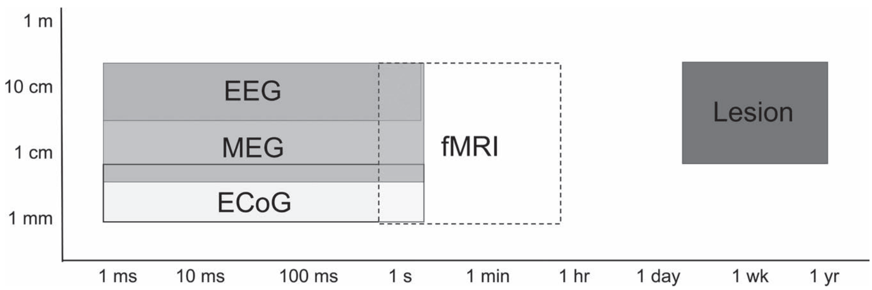


Figure 8.1 Measurement techniques discussed in this chapter are depicted in terms of their spatial and temporal resolution. Scalp EEG, scalp electroencephalogram; MEG, magnetoencephalogram; iEEG, intracranial electroencephalogram; fMRI, functional magnetic resonance imaging. Each box indicates the spatial/temporal resolution for each method. These boxes reflect a range of possible values, which can vary on the basis of particular equipment or methods used.

One benefit neuroimaging approaches have over a neuropsychological approach is that they allow for the study of mental imagery and music in the healthy brain. Additional benefits of fMRI are the high spatial resolution, relative accessibility, and a large community of fMRI researchers. For the purposes of the present topic (music and mental imagery), one limitation of fMRI is that the environment is very loud. Of course, this does not eliminate the possibility of conducting auditory experiments using fMRI. To circumvent the issue of scanner noise, some researchers have adopted a "sparse sampling" technique, where MRI measurements are taken after the stimuli are presented in silence (e.g., Norman-Haignere et al., 2015). Other limitations for MRI research are the costs associated with scanning (typically hundreds of dollars per hour) and the accessibility of such methods for studying certain populations (e.g., movement is quite disruptive, making fMRI a challenging technique to use with children).

Examples of fMRI Studies of Music and Imagery

A frequent question in the neuroscience of musical imagery is the extent to which music perception (i.e., actual listening) and musical imagery (i.e., imagined listening) are similar in their neural representation. In one experiment examining this question, participants were shown the lyrics of familiar musical tunes, and either heard the tune along with the lyrics or were asked to imagine the tune (Herholz et al., 2012). Researchers found that listening to tunes activated a wide range of auditory regions, including primary auditory cortex. Imagining the melodies also activated regions in auditory association cortices, but not in primary auditory cortex. This suggests that auditory imagery engages some of the same neural structures as imagery, but not primary sensory regions. Another interesting feature of this study is that they looked at how individual differences in auditory imagery ability (as measured by the Bucknell Auditory Imagery Scale) related to neural activity during imaging melodies. The authors found that participants with greater overall imagery ability showed more neural activity in auditory regions during musical imagery.

While most fMRI work investigating music and imagery has focused on identifying the neural overlap between music perception and musical imagery, musical (or auditory) imagery is not the *only* type of mental imagery associated with music (see chapters in Part I of this volume). For example, visual imagery may be evoked during music listening (see Taruffi & Küssner, this volume), perhaps associated with a memory induced by the music or just in response to the music itself (e.g., a listener may hear a song that reminds them of rushing water, evoking a mental image of a waterfall). In this case, the listener is perceiving music, but the imagery is a visual mental image. One recent study combined the use of fMRI with a pharmacological intervention to investigate the neural correlates of music-induced visual imagery. Researchers sought to identify how the drug LSD modulated neural responses to music-induced visual imagery (Kaelen et al., 2016). Participants underwent fMRI scanning after receiving intravenous LSD or placebo. During the scanning, participants either listened to music or silence, and afterwards, they rated the degree of visual imagery experienced. The authors found a significant interaction between drug and listening condition: Participants who received LSD showed increased neural coupling between the parahippocampal cortex and primary visual cortex, but only during music listening. They also found a positive correlation between imagery ratings and parahippocampal-visual cortex connectivity. The authors interpret these results to suggest that LSD and music, when combined, enhance vivid visual imagery.

Electroencephalography

Electroencephalography (EEG) is a measurement technique for recording the electrical activity of groups of neurons. Given that this technique measures electrical activity (rather than changes in blood oxygenation, as in fMRI), the temporal resolution is quite precise (on the order of milliseconds, rather than seconds; see Figure 8.1). There are two variations of EEG, which greatly differ in their spatial resolution: Scalp EEG is recorded from a series of electrodes placed on the scalp (typically ranging from 16 to 128 electrodes on a cap the participant wears). In this case, the neural signal diffuses through the skull and scalp, making it challenging to localize the activity to specific brain regions. In contrast, during intracranial EEG (iEEG), a grid of electrodes is placed directly on the cortical surface (in the case of electrocorticography or ECoG), or depth electrodes are placed in subcortical regions (called stereotactic-EEG). Therefore, this method has the advantage of very precise spatial resolution (on the order of millimetres) as well as precise temporal resolution (on the order of milliseconds, as with scalp EEG). In contrast to scalp EEG, which can be used to study the healthy brain, iEEG is used in patients with medically intractable epilepsy who are undergoing intracranial monitoring prior to resection of epileptic tissue. For more detailed information on iEEG, see Parvizi and Kastner (2018).

When conducting EEG experiments (both scalp and intracranial), participants typically perform several hundreds of trials. One approach to analysing EEG data is to average the neural responses over multiple trials to extract “event-related potentials” (ERPs). The ERP technique has been used frequently in studies of music perception and cognition (see the “Early Right Anterior Negativity;” Koelsch et al., 2005) and provides precise information about the timing of neural responses. In contrast to ERPs, an alternative approach is to look at the oscillatory activity in certain frequency bands (e.g., alpha, beta, and gamma).

The key advantage of both scalp and intracranial EEG is the precise temporal resolution. Other advantages of scalp EEG are that it is relatively inexpensive and silent, so it would not disrupt auditory stimuli (as compared to fMRI). Measuring scalp EEG does not necessarily constrain the participant in the same way as larger machines such as MRI. Recently, mobile EEG equipment has been developed to use in more naturalistic settings such as concert halls or art museums (Herrera-Arcos et al., 2017; Zamm et al., 2017). The primary limitation of scalp EEG is its relatively poor spatial resolution. Source-reconstruction methods can be used, but because the EEG signal is diffused through the skull, it is difficult to precisely identify the source in the brain.

Intracranial EEG has the distinct advantage of being precise in both temporal and spatial resolution. However, there are some obvious limitations: For ethical reasons, iEEG can only be used in patients requiring it for clinical diagnosis or treatment. Therefore, subject numbers are small, and the location that electrodes are placed must be determined on a clinical basis—this means that certain brain regions are more likely to be studied than others. Additionally, any results should be taken with the caveat that they are coming from an "atypical" brain. Finally, iEEG is also only done at certain hospitals and requires the cooperation of neurosurgeons and other clinicians, and is therefore less accessible than other methods.

Examples of EEG Studies of Music and Imagery

One EEG study sought to investigate differences in neural activity in the alpha frequency band between perceived and imagined music (Schaefer et al., 2011). In this experiment, participants heard a musical phrase (the *perception* stage) followed by silence, during which they were to imagine the phrase in their heads (the *imagery* stage). Stimuli consisted of approximately 3-second excerpts of a classical and a pop piece of music. The authors found a stronger alpha band response in occipito-parietal regions during imagery versus perception. They interpret this result with regard to other work indicating increased alpha power during auditory working memory tasks, and suggest that musical imagery engages working memory processes.

A recent ECoG study also investigated similarities and differences in the neural responses to music perception and imagery (Martin et al., 2018). In this study, a single patient undergoing ECoG for medically intractable epilepsy completed two experimental conditions. In the first condition, he played two pieces of music on a piano keyboard with auditory feedback. In the second, he played the same piece but did not receive any auditory feedback. During the second condition, he was asked to imagine what it would sound like while he was playing. This study found that activity in auditory-related regions showed substantial (but not total) overlap between the perception and imagery conditions, suggesting that similar networks involved in both music perception and imagery. Additionally, due to the high temporal and spatial resolution of ECoG, the authors were able to reconstruct the auditory spectrogram from high gamma signals recorded from the cortical surface (high gamma is a higher frequency that can be measured using scalp EEG). They found that the accuracy of these reconstructions did not differ between the perceived and imagined conditions.

Magnetoencephalography

In addition to generating an electrical signal, neural activity generates a corresponding magnetic field. Magnetoencephalography (MEG) is a technique used to measure this magnetic field. Therefore, like EEG, MEG has relatively high temporal resolution, and a researcher can analyse the MEG signal using either an evoked-potential or a time–frequency approach. Experimental design of MEG experiments follows the same principles of EEG and fMRI (e.g., well-designed tasks with sufficient repeated trials).

One advantage of MEG is the precise temporal resolution, which is similar to EEG. In contrast to EEG, MEG provides relatively good spatial resolution. Because magnetic fields do not diffuse as greatly through the surface of the skull as electrical signals, MEG signals are more easily localized than those measured using EEG. MEG therefore has many advantages, but it has not been as readily adopted as EEG or fMRI for several reasons: For one, MEG equipment is quite costly (on the order of millions of dollars), although these costs are similar to those of MRI. One major difference between MRI and MEG is that standard MRI is very commonly used in clinical settings, making it a more accessible method. Most researchers at large universities likely have some sort of access to MRI facilities. Currently, MEG does not have the same type of clinical applications, and is typically used primarily for research. This makes it relatively inaccessible, unless a researcher is located at one of the few hundred hospitals or universities with access to MEG equipment.

Examples of MEG Studies of Music and Imagery

One study used MEG to investigate the oscillatory activity underlying musical hallucinations in a healthy adult with no history of neurological disorder (Kumar et al., 2014). In the experimental task, the participant heard 30 seconds of classical music to "mask" her hallucinations, followed by a return in the typical hallucination pattern after the music stopped. Throughout the experiment, the participant rated the strength of her hallucinations. The authors identified that gamma power increases, localized to the left hemisphere superior temporal gyrus as well as motor and premotor areas, were associated with stronger musical hallucinations. These data seem to complement the data from lesion studies, suggesting that left hemisphere temporal lobe structures are implicated in musical hallucinations.

A recent MEG study sought to investigate neural signatures of imagery for musical pitch, as isolated from rhythm and other temporal aspects of music (Gelding et al., 2019). Participants completed the Pitch Imagery Arrow Task (see also Gelding et al., this volume), which consisted of listening to a major scale, followed by either the tonic or dominant note in the scale. Participants then saw an arrow pointing up or down to indicate which note they should imagine (above or below the preceding note). For perception trials, participants heard this note instead of imagining it. The authors found increased magnitude of beta activity during imagery versus perception. The authors suggest that this indicates that not only is beta activity involved in the timing and rhythmic aspects of musical perception and imagery (as suggested in prior research), but it is also important for imagery of musical pitch.

Summary and Conclusions

This chapter reviewed several commonly used neuroscience methods to study music and mental imagery. However, there are other neuroscientific methods and analysis techniques that were not included in the present work. For example, while only functional MRI was discussed here, researchers have also used structural MRI to study music and imagery (Farrugia et al., 2015). Other neuroimaging methods or techniques, such as Positron Emission Tomography or Diffusion Tensor Imaging (an MRI-based technique for assessing white matter tracts), have not frequently been used to investigate this topic. In reviewing the most often used neuroscience techniques in the study of music and imagery, the goal was to provide an introduction to a variety of methods and offer some practical guidelines for choosing which methods to implement in future studies on this topic.

As mentioned at the outset of the chapter, neuroscience approaches are likely of little value to the study of music and mental imagery without accompanying behavioural and/or self-report measures (see Gelding et al., this volume; Hubbard, this volume). As has been suggested before, "neuroscience needs behavior" (Krakauer et al., 2017). Hopefully readers interpret this chapter in the greater context of this book, particularly the chapters describing other measurement techniques for studying music and mental imagery. When concluding this chapter, it is important to mention a key limitation of neuroimaging approaches overall: Interpreting the meaning of brain activity is challenging. One common finding in neuroimaging work on music and mental imagery is that mental imagery is associated with activity in sensory-specific cortices, without actual sensory stimulation. This may lead the researcher to make "reverse inferences," a common pitfall of neuroimaging research (Poldrack, 2011b). For example, if researchers observe increased visual cortex activity during music listening, they may infer that the listener was engaging in active visual imagery. This is a common approach to interpreting neuroimaging data, but unless the behaviour of interest is actually being measured, this reverse inference (inferring behaviour based on brain activity) is entirely speculative (see also Ofner & Stober, this volume).

Such interpretations would therefore be strengthened by corresponding behavioural and self-report data: Should a participant report that they were engaging in visual imagery, *and* show increased activity in visual cortices, this would provide a stronger case than either behaviour (and/or self-report) or neuroimaging data alone. Overall, neuroscience methods have added to our understanding of music and mental imagery, suggesting that vivid mental imagery is similar to perception, in that it activates sensory-specific cortices. However, for a greater understanding of music-evoked imagery, it is important to approach this interdisciplinary problem using complementary methods from neuroscience and other measurement techniques.

References

Belfi, A. M., Karlan, B., & Tranel, D. (2018). Damage to the medial prefrontal cortex impairs music-evoked autobiographical memories. *Psychomusicology: Music, Mind, and Brain*, *28*(4), 201. https://doi.org/10.1037/pmu0000222

Belfi, A. M., & Tranel, D. (2014). Impaired naming of famous musical melodies is associated with left temporal polar damage. *Neuropsychology*, *28*(3), 429–435. https://doi.org/10.1037/neu0000051

Bernardini, F., Attademo, L., Blackmon, K., & Devinsky, O. (2017). Musical hallucinations: A brief review of functional neuroimaging findings. *CNS Spectrums*, *22*(5), 397–403. https://doi.org/10.1017/S1092852916000870

Calabrò, R. S., Baglieri, A., Ferlazzo, E., Passari, S., Marino, S., & Bramanti, P. (2012). Neurofunctional assessment in a stroke patient with musical hallucinations. *Neurocase*, *18*(6), 514–520. https://doi.org/10.1080/13554794.2011.633530

Dellacherie, D., Bigand, E., Molin, P., Baulac, M., & Samson, S. (2011). Multidimensional scaling of emotional responses to music in patients with temporal lobe resection. *Cortex*, *47*(9), 1107–1115. https://doi.org/10.1016/j.cortex.2011.05.007

Farrugia, N., Jakubowski, K., Cusack, R., & Stewart, L. (2015). Tunes stuck in your brain: The frequency and affective evaluation of involuntary musical imagery correlate with cortical structure. *Consciousness and Cognition*, *35*, 66–77. https://doi.org/10.1016/j.concog.2015.04.020

Gelding, R. W., Thompson, W. F., & Johnson, B. W. (2019). Musical imagery depends upon coordination of auditory and sensorimotor brain activity. *Scientific Reports*, *9*(1), 1–13. https://doi.org/10.1038/s41598-019-53260-9

Golden, E. C., & Josephs, K. A. (2015). Minds on replay: Musical hallucinations and their relationship to neurological disease. *Brain*, *138*(12), 3793–3802. https://doi.org/10.1093/brain/awv286

Gosselin, N., Peretz, I., Hasboun, D., Baulac, M., & Samson, S. (2011). Impaired recognition of musical emotions and facial expressions following anteromedial temporal lobe excision. *Cortex*, *47*(9), 1116–1125. https://doi.org/10.1016/j.cortex.2011.05.012

Herholz, S. C., Halpern, A. R., & Zatorre, R. J. (2012). Neuronal correlates of perception, imagery, and memory for familiar tunes. *Journal of Cognitive Neuroscience*, *24*(6), 1382–1397. https://doi.org/10.1162/jocn_a_00216

Herrera-Arcos, G., Tamez-Duque, J., Acosta-De-Anda, E. Y., Kwan-Loo, K., de-Alba, M., Tamez-Duque, U., . . . Soto, R. (2017). Modulation of neural activity during guided viewing of visual art. *Frontiers in Human Neuroscience*, *11*, 581. https://doi.org/10.3389/fnhum.2017.00581

Huettel, S. A., Song, A. W., & McCarthy, G. (2014). *Functional magnetic resonance imaging* (3rd ed.). Oxford: Oxford University Press.

Kaelen, M., Roseman, L., Kahan, J., Santos-Ribeiro, A., Orban, C., Lorenz, R., . . . Carhart-Harris, R. (2016). LSD modulates music-induced imagery via changes in parahippocampal connectivity. *European Neuropsychopharmacology*, *26*(7), 1099–1109. https://doi.org/10.1016/j.euroneuro.2016.03.018

Koelsch, S., Gunter, T. C., Wittfoth, M., & Sammler, D. (2005). Interaction between syntax processing in language and in music: an ERP Study. *Journal of Cognitive Neuroscience*, *17*(10), 1565–1577. https://doi.org/10.1162/089892905774597290

Krakauer, J. W., Ghazanfar, A. A., Gomez-Marin, A., MacIver, M. A., & Poeppel, D. (2017). Neuroscience needs behavior: Correcting a reductionist bias. *Neuron*, *93*(3), 480–490. https://doi.org/10.1016/j.neuron.2016.12.041

Kumar, S., Sedley, W., Barnes, G. R., Teki, S., Friston, K. J., & Griffiths, T. D. (2014). A brain basis for musical hallucinations. *Cortex*, *52*, 86–97. https://doi.org/10.1016/j.cortex.2013.12.002

Martin, S., Mikutta, C., Leonard, M. K., Hungate, D., Koelsch, S., Shamma, S., . . . Pasley, B. N. (2018). Neural encoding of auditory features during music perception and imagery. *Cerebral Cortex*, *28*(12), 4222–4233. https://doi.org/10.1093/cercor/bhx277

Moseley, P., Alderson-Day, B., Kumar, S., & Fernyhough, C. (2018). Musical hallucinations, musical imagery, and earworms: A new phenomenological survey. *Consciousness and Cognition*, *65*, 83–94. https://doi.org/10.1016/j.concog.2018.07.009

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*(6), 1281–1296. https://doi.org/10.1016/j.neuron.2015.11.035

Parvizi, J., & Kastner, S. (2018). Promises and limitations of human intracranial electroencephalography. *Nature Neuroscience*, *21*(4), 474–483. https://doi.org/10.1038/s41593-018-0108-2

Pearson, J., Naselaris, T., Holmes, E. A., & Kosslyn, S. M. (2015). Mental imagery: Functional mechanisms and clinical applications. *Trends in Cognitive Sciences*, *19*(10), 590–602. https://doi.org/10.1016/j.tics.2015.08.003

Poldrack, R. A. (2011a). *Handbook of functional MRI data analysis* (1st ed.). Cambridge: Cambridge University Press.

Poldrack, R. A. (2011b). Inferring mental states from neuroimaging data: From reverse inference to large-scale decoding. *Neuron*, *72*(5), 692–697. https://doi.org/10.1016/j.neuron.2011.11.001

Samson, S., Zatorre, R. J., & Ramsay, J. O. (2002). Deficits of musical timbre perception after unilateral temporal-lobe lesion revealed with multidimensional scaling. *Brain*, *125*(3), 511–523. https://doi.org/10.1093/brain/awf051

Schaefer, R. S., Vlek, R. J., & Desain, P. (2011). Music perception and imagery in EEG: Alpha band effects of task and stimulus. *International Journal of Psychophysiology*, *82*(3), 254–259. https://doi.org/10.1016/j.ijpsycho.2011.09.007

Zamm, A., Palmer, C., Bauer, A.-K. R., Bleichner, M. G., Demos, A. P., & Debener, S. (2017). Synchronizing MIDI and wireless EEG measurements during natural piano performance. *Brain Research*, *1716*, 27–38. https://doi.org/10.1016/j.brainres.2017.07.001

9 Deep Neural Networks and Auditory Imagery

André Ofner and Sebastian Stober

Decoding Auditory Information From Neuroimaging Data

Two different strategies have been pursued to explore the links between cognitive processes and brain activity in the neuroimaging domain. The typical procedure in forward inference is the manipulation of a subject's psychological state followed by a calculation of the probabilities of observing specific brain signals given this state, that is, a process from psychological state to expected brain activity (Geuter et al., 2017). In contrast, reverse inference is used to reason in the other direction, that is, from brain activity to cognitive processes. Reverse inference from brain activity is a complex endeavour, especially as it requires knowledge about the information the analysed brain signals actually can account for. This is especially problematic when reasoning from specific regions of the brain, as their activation could be the result of reused activations within different cognitive processes. In reverse inference, "decoding" approaches typically model the space of brain signal as multivariate and the space of cognitive processes as univariate. In contrast, "encoding" approaches typically model a univariate brain space and a multivariate space of cognitive processes. Some methods, such as canonical correlation, consider both neural and psychological space as multivariate. These multivariate methods allow to perform inference with increased accuracy, especially as they consider spatio-temporal relations across brain regions. In recent years, the use of machine-learning algorithms to decode information from brain activity has gained widespread attention and several studies have started to apply them to the analysis of audio- and imagery-related neural processes. Brain activity decoding in the presence of auditory imagery promises insights into the contributing cognitive functions during music listening and imagination while also shining a light on the multimodal mapping to (imagined) sensory inputs.

A general consensus in these studies is that a listener's brain shows responses that are correlated with presented auditory stimuli and that the relationship between brain response and stimulus can be exploited to classify or reconstruct auditory stimuli. Examples for such correlations can be found in the modulation of neural oscillation magnitude and frequency by perceived tempo, rhythm, and accents (Cirelli et al., 2014; Nozaradan et al., 2012). Intracranial recordings show precise phase-locking to click train stimuli and frequency-following responses for

DOI: 10.4324/9780429330070-12

speech stimuli (Nourski et al., 2009; Krishnan & Plack, 2011). Auditory event-related potentials (ERPs) refer to repeatable and distinguishable neural responses to auditory events, such as onsets or changes in pitch (Schaefer et al., 2009). ERPs can be related to fine-grained aspects of audio, reflecting even differences in timbre or harmonics (Shahin et al., 2005). Neural responses are modulated by individual aspects such as expertise or attention (Treder et al., 2014). This means that methods aiming at reverse inference benefit from modelling both the subject's particular behaviour as well as taking into account the structure of auditory stimuli and environment. The idea of mapping auditory features to brain signals has found traction in a variety of studies applying machine learning to reconstruct aspects such as the loudness envelope, tempo, or pitch (Sternin et al., 2015; Stober et al., 2015). These studies are focused primarily on audio perception and show promising results, even if the quality of stimulus decoding is rather unsatisfactory. O'Sullivan et al. (2015) describe a reconstruction of speech stimulus envelopes directly from recorded EEG signal based on the correlations between EEG channels and the stimulus envelope. However, as described by Stober (2017), the application of the same approach to more complex musical signals leads to poor results and demonstrates the low correlation between recorded signal and auditory stimulus, which is typical for non-invasive imagining methods. A common way to deal with such low correlation is to average across a large number of trials and focus on relatively simple stimuli, a technique particularly popular in ERP-based experiments (Woodman, 2010). Neural activity recording with electroencephalography (EEG) or functional magnetic resonance imaging (fMRI) provides relatively accurate and inexpensive access to brain signal, making it particularly interesting for machine learning. While EEG has high temporal resolution, the spatial accuracy is inferior to the three-dimensional fMRI signal. Multimodal recording techniques like simultaneous EEG–fMRI capture multiple views on brain activity in spatially and temporally overlapping regions and allow more fine-grained analysis of the signal (Huster et al., 2012). For more detailed information on imaging techniques, see Belfi (this volume).

From Auditory Perception Towards Imagery Decoding

Most aforementioned studies focus primarily on the perception of audio and music, in contrast to music imagination or combinations of both conditions (Stober et al., 2015). However, multiple studies using EEG and MRI indicate that large parts of the neural processes underlying music perception can also be observed in absence of the actual stimulus (Herholz et al., 2012; Schaefer, 2014). The literature on auditory imagery in the context of decoding neuroimaging data tends to deal with imagery and imagination as interchangeable terms. Often times, authors are interested in the analysis or reconstruction of mental imagery that takes place in the process of active imagination. In these cases, this lack of separation is not too problematic. However, there might be cases where a more fine-grained differentiation is necessary, such as in auditory imagery during memory recall or active imagination.

Here we refer to auditory imagery as the process of experiencing auditory information without it being present at the senses. Furthermore, we treat auditory imagination as a process that can employ mental auditory images, such as when actively imagining a specific song. Auditory imagery generally retains the structure of actual auditory content, which is shaped by aspects such as tempo and pitch. Individual aspects of the auditory signal might be more or less accurately retained when recalling previously heard stimuli, for example, when the timbre but not the exact tempo is preserved.

Auditory imagery has been shown to be interacting with other imagery modalities, such as motor imagery (see Godøy, this volume). For example, being skilled in motor and auditory imagery has positive impact on the learning performance in the respective other domain (Brown et al., 2013). This indicates that auditory imagery can be seen as a process that reactivates the neural pathways underlying auditory perception and recreates an internalized sensory experience. This implies that features extracted from brain signal in perception across different modalities should be useful for guiding the decoding of brain activity during imagination. Still, only a relatively small number of studies have tackled the problem of classifying brain states in auditory imagery or reconstructing the imagined stimuli directly (González et al., 2019; Ofner & Stober, 2018).

Recently, González et al. (2019) demonstrated the possibility to use auditory imagery decoding in a brain–computer interface setup using EEG hardware. The study focused on the recognition of white noise imagery and reported 93% accuracy. According to the authors, these results open up the possibility to use auditory imagery as a complementary approach to motor-imagery interfaces. This aspect is especially interesting given the close connection between motor and auditory imagery. Ofner and Stober (2018) addressed the direct reconstruction of imagined stimuli from EEG. They trained a generative deep neural network to map both EEG and audio signal to a shared representation. The trained model was subsequently applied to reconstruct Mel spectrograms of the same sentences during imagination. In the experiment, the imagination directly followed the sentence perception, with additional temporal guidance using a constant metronome click. The reconstructed stimuli showed alignment in tempo and timbre with the original stimuli both in the perception and in imagination condition. More work has been done in the area of motor imagery decoding (Aggarwal & Chugh, 2019). Combining models trained on motor imagery with those focusing on auditory information could increase their accuracy and interpretability. Another direction of research could explore the interdependencies between visual and auditory imagery. At the time of writing, there is still a lack of machine learning studies bridging between different imagery domains and exploring these interdependencies in greater depth.

Deep Neural Networks and Auditory Imagery

Artificial intelligence is a thriving field that has generated a large variety of new research directions and applications for intelligent software, such as automatic speech and image recognition and autonomous driving. Different strategies have been explored on the quest towards developing intelligent machines. Prominently, there are projects that are focused on relatively formal and abstract tasks, most of which can be solved by using formal rules or large knowledge bases. Solving formal problems is relatively easy for a computer and more difficult for humans. Interestingly, tasks that seem much simpler for humans, such as recognizing objects in everyday life, pose a significantly bigger challenge to machine intelligence. The underlying problem of extracting meaningful knowledge from sensory information and behaving in an intelligent way in complex environments has become one of the most important directions of investigation in current artificial intelligence research. Here, the information that is extracted is less easy to formalize and is often times subjectively or intuitively processed in humans. The field of research that aims at extracting such knowledge from raw data is called machine learning (ML).

A central aspect of machine learning algorithms is their capacity to represent information. Many ML algorithms, such as logistic regression or naïve Bayes, work on representations that are defined before solving a particular task. Within the field of ML, the set of representations that are available to the algorithm to make inference are called features. For many applications in neuroimaging, it might be sufficient to provide specifically chosen features, such as the activation of a specific brain region or frequency bands. Often times, however, it is difficult to find the right set of predefined features, especially when dealing with data as complex as in the neuroimaging domain. These considerations have led to the subfield of representation learning in ML. Representation learning algorithms no longer project points from a predetermined feature space to outcomes. Instead, they allow to learn the representation itself. Such representation learning methods can extract complex sets of features, often times much more complex and predictive than hand-crafted features. Many representation learning algorithms furthermore allow to find sets of features without human supervision, allowing to tackle large and diverse amounts of data, an aspect that is of special importance in the application to neuroimaging data. An important example for such representation learning algorithms are autoencoders, which will be dealt with in more detail later in this chapter. Autoencoders learn representations of raw data by converting inputs into an internal representation, which in turn is decoded back into the original input. This way, autoencoders are trained to preserve as much information as possible, even if the internal representation is reduced in size. An important aspect underlying good representations are their capacity to explain and disentangle the factors of variation in the data. In the example of EEG signal, a representation learning algorithm might learn to separate factors such as age, gender, or attention in order to explain the processed data. Similarly, an algorithm working on images might learn that an object's colour is affected by the time of day. An important problem that arises here is the aspect of extracting high-level features from raw data. This challenge has been addressed with neural networks (NNs) and particularly the class of deep neural networks (DNNs), which are based on the idea of expressing complex representations as weighted combinations of simpler representations. These computational models excel at statistical pattern recognition, especially when trained on large datasets, often times containing millions of (labelled) data points. A vast selection of different network architectures has been developed, often tailored to deliver high accuracy on a particular task (LeCun et al., 2015; Shrestha & Mahmood, 2019). Common to deep learning models is that they are built from multiple layers that learn representations with increasing level abstraction, as information propagates towards deeper, hidden neural layers. DNNs are usually trained using the gradient descent algorithm. During learning, the internal parameters of the network are changed based on the error made on a given task. To achieve this, the gradients of a loss function, which captures the magnitude of the error, are computed with respect to the network weights. Through parameter modification, the network's representational activity is changed, as each layer computes representations based on those arising in previous layers. Figure 9.1 (a) provides a schematic overview of the typical layers found in DNNs that transform an input to predicted outputs.

Several aspects underlying this approach to machine learning bear resemblance to neural architectures found in the biological brain. For example, deep convolutional neural networks (CNNs) show connectivity similar to the organization of neurons in the visual cortex, where individual neurons respond only to activation within their respective receptive fields. This CNN architecture is in turn a refinement of the "neocognitron" published in 1980 by Fukushima (Fukushima, 1980). In CNNs, these receptive fields are efficiently computed across input dimensions, and increasingly complex representations are learned through combining features in overlapping receptive fields. This allows CNNs to compute abstract representations with receptive fields that can cover the entire input. DNNs have significantly improved state-of-the-art results in a wide range of fields, such as visual object detection or speech recognition and synthesis (LeCun et al., 2015). CNNs are particularly suited for spatial processing, as they excel at aggregating increasingly abstract features and their spatial relations. These features can then be used for input reconstruction, classification, or regression tasks. Next to CNNs, recurrent neural networks (RNNs) are an important class of deep neural networks. In contrast to CNNs, RNNs feature recurrent or feedback connections between internal states in the network, making them a good choice for processing sequential data. During training, these recurrent connections between internal states of the model are "unfolded" in time. *Unfolding* refers to the process of making the timesteps in the network explicit, resulting in a structure that allows to update network weights analogous to regular DNNs. RNNs can be used for sequence modelling and provide access to memory states, which can be, for example, used to predict future samples of a given signal. Figure 9.1 (b & c) provides an overview of the key structural differences between CNNs and RNNs and shows an example for RNN unfolding with three timesteps. Both CNN- and RNN-based approaches have found use in the analysis of neural activity underlying auditory perception and imagination. They have demonstrated improved results for tasks such as tempo estimation or classification of imagined musical pieces (Moinnereau et al., 2018; Stober, 2017). However, often times neuroimaging datasets on auditory data do not contain enough data to train large models. Typical datasets in the domain

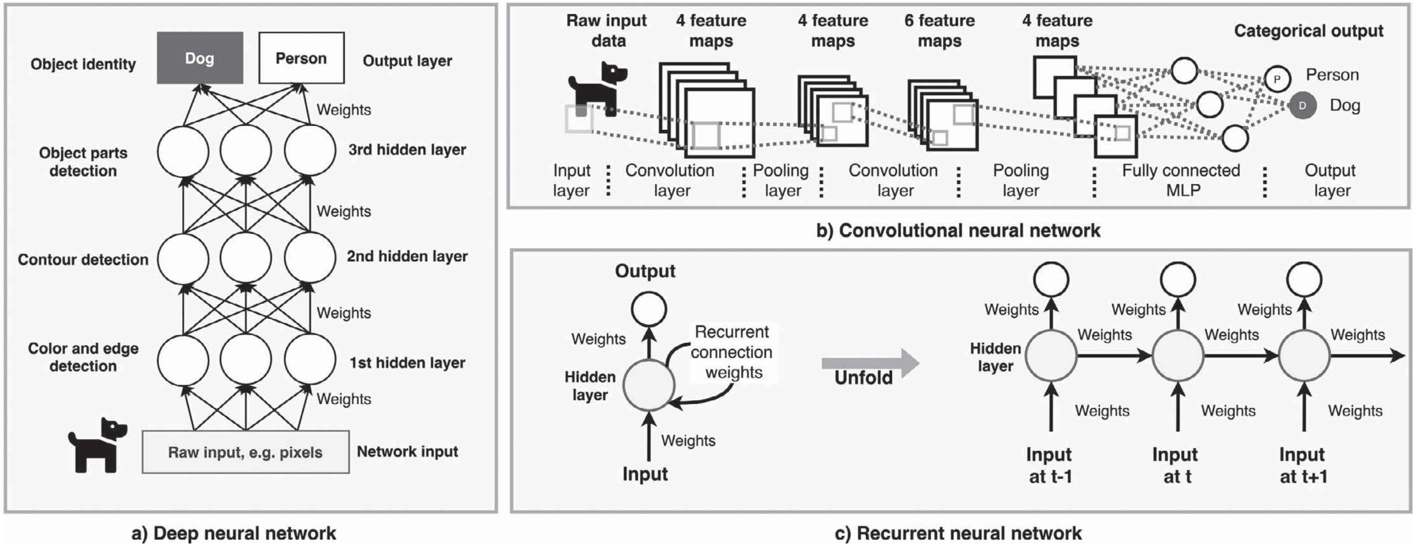


Figure 9.1 Comparison of dense, convolutional, and recurrent layers in deep neural networks.

contain 20 or fewer subjects and cover only short excerpts of musical pieces or imagined stimuli. Only very few of the available datasets contain larger amounts of subjects or provide more than one hour of recording time per subject. These more extensive datasets typically cover the perception of full songs or movies, or are recorded in the context of auditory brain–computer interfacing (Daly et al., 2020, Losorelli et al., 2017). This shortage could be counteracted with more and longer recording sessions and detailed metadata.

Auditory Imagery Decoding With Deep Generative Models

Recently, various deep generative models have been introduced that combine the representational capacity of DNNs with unsupervised learning and stochastic inference. Popular examples are autoencoders (AE) and variational autoencoders (VAE). AEs and VAEs learn to reconstruct data based on an internal, latent representation. Generative models enable unsupervised learning, that is, the automatic extraction of meaningful representations from data without annotations through the process of encoding and decoding from a low-dimensional latent representation. This is in contrast to non-generative approaches, such as CNNs, which lack such an internal representation that lends the ability to generate new data. More precisely, the idea of autoencoding refers to the process of reconstructing an input from an internal representation using a so-called decoder network. This encoding is computed by a separate network, the encoder, or (variational) inference network. While autoencoders in general have been in use for decades, the combination of generative modelling and deep learning is a more recent development. The dimensionality of the learned representation typically is significantly smaller than the input data, often times forcing the network to ignore noise in the data. VAEs allow to perform Bayesian inference (see also Vroegh, this volume) and feature a stochastic process (Kingma & Welling, 2013). In comparison, AEs are entirely deterministic models. While both network types are able to compute compressed latent representations of the input data, VAEs allow to control the latent distributions and are known to disentangle factors that generate the data. Typically, this is done by modelling Gaussian priors on the latent variables, which pushes different factors into different regions of the latent space. A popular and relatively simple example for this process is the separation of the handwriting style from the written character. However, the same idea can also be applied to representing generating factors of neural activity or auditory information. For example, aspects of a musical piece can be represented by the latent factors—loudness and pitch. Variants of the VAE architecture have been used to encode brain states and perceived and imagined stimuli into a shared latent representation (Ofner & Stober, 2018). Crucially, as the employed model is generative, the learned representation can be used to decode a distribution of likely sensory stimuli at any point in time. Unfortunately, training and analysing generative models such as VAEs is not straightforward. For example, the quality of the generated stimulus reconstructions is generally measured by a simple mathematical distance between predicted and presented stimulus during perception. This distance, however, might not capture musically significant aspects of the stimulus. For example, a simple temporal shift of an otherwise entirely correct prediction can already lead to a large mathematical error. For predicted imagined stimuli, the problem is even more severe: During imagination, the provided stimulus can only be an estimation of the actually imagined one. Nevertheless, generative models are a useful tool for investigations of auditory imagery, especially when used in real time or in interactive settings, such as in brain–computer interfaces (BCIs). In such settings, generative models provide the possibility to generate meaningful stimuli which could be employed as feedback for the subject and to refine the model itself. One could imagine, for example, audible feedback being generated in real time and the immediate response of the subject being used to further modify the predicted signal. Future studies could improve the accuracy of generative models with more extensive datasets: Providing multimodal neuroimaging signals can allow generative models to learn more meaningful representations. Here, multimodality can refer either to multimodal imaging techniques or to multimodal neuronal systems, such as when jointly analysing activities in auditory or motor systems. Multimodal imaging can be integrated with auditory decoding, for example, by reconstructing both EEG and fMRI signals along with the predicted stimulus. Using multiple imaging sources deliver views on the same neural activation patterns which capture the signal at different spatial and temporal resolutions. For example, temporal processes that are detected in EEG could be used to predict the more spatially accurate fMRI signal and vice versa. However, using multiple views on neural activity also brings new challenges, such as the induced artefacts and delays between the recorded signal sources (Abreu et al., 2018). A possible approach to the second type of multimodality is the prediction of recorded physiological or video signal capturing aspects such as the subject's facial or motor activity simultaneously with the neural signature. This would allow to learn representations that encode similarities and differences, for example, between visual and auditory imagery, as image and audio could now be decoded simultaneously. Finally, datasets that provide a direct guidance for the decoding of imagery can help to create improved models. This could be achieved by providing the score to a musical work or by assisting the imagination process with a metronome. This way, parts of the uncertainty can be reduced, aiding both network training and the measurability of its accuracy.

Future Directions in Machine Learning-Based Auditory Imagery Research

While generative models provide powerful tools for analysis and prediction, they could further benefit from a more biologically plausible structure. One promising direction of research with respect to auditory imagery is the fusion of deep learning techniques with predictive coding (PC) or active inference, also referred to as free energy principle (FEP; Aitchison & Lengyel, 2017; Pezzulo et al., 2015; Rao & Ballard, 1999). Predictive coding delivers a computational description of neural function that can be interpreted as a Bayesian view on brain function. In predictive coding, the brain generates (Bayesian) predictions about sensory inputs and corrects its representation of the world based on the prediction errors. In line with the more general FEP, this process is hypothesized to explain the internally generated sensory experience as observed in sleep or mental imagery, where internal models are tested and refined. Various deep learning models inspired by PC and the FEP exist, but there is still a lack of studies that apply them to brain signal decoding (Lotter et al., 2016; Ofner & Stober, 2019). Multiple studies have explored the modulation of brain activity based on prediction errors in the auditory and visual domain. For example, auditory mismatch negativity brain potentials have been explained with predictive coding, where auditory attention is modulated by prediction errors (Auksztulewicz & Friston, 2015; Stefanics et al., 2018). More specifically, a hierarchy of precision-weighted prediction errors modulate auditory attention in such a way that event-related responses to unexpected events are enhanced, and expected events evoke smaller responses. Predictive models based on PC thus can take into account individual aspects, such as the immediate or long-term history of a subject's emotional response or memory recall in order to explain a particular neural response or lack thereof. Future studies could explore, for example, changes in auditory surprise (or prediction error) with and without imagery. Using such PC- or FEP-based models, insights about predictions and resulting prediction errors could be generated in the human and artificial models simultaneously, for example, by exposing both to the same set of stimuli. Importantly, this signifies a step away from directly correlating brain activity and stimulus and instead modelling the complex relationship between a subject's individual behaviour and the characteristics found in the (imagined) auditory stimulus. Methodologically speaking, this could be done individually for each new stimulus or similar to the typical procedure in ERP studies, by averaging across many instances of predictions. The detectable processes in brain signals could then be compared quantitatively and qualitatively with the predictions generated by the model. This way, exploring functional mechanisms in auditory imagery could be combined with the capacity of deep neural networks to process large amounts of complex data.

References

Abreu, R., Leal, A., & Figueiredo, P. (2018). EEG-informed fMRI: a review of data analysis methods. *Frontiers in Human Neuroscience*, *12*, 29. https://doi.org/10.3389/fnhum.2018.00029

Aggarwal, S., & Chugh, N. (2019). Signal processing techniques for motor imagery brain computer interface: A review. *Array*, *1*, 100003. https://doi.org/10.1016/j.array.2019.100003

Aitchison, L., & Lengyel, M. (2017). With or without you: Predictive coding and Bayesian inference in the brain. *Current Opinion in Neurobiology*, *46*, 219–227. https://doi.org/10.1016/j.conb.2017.08.010

Auksztulewicz, R., & Friston, K. (2015). Attentional enhancement of auditory mismatch responses: A DCM/MEG study. *Cerebral Cortex*, *25*(11), 4273–4283. https://doi.org/10.1093/cercor/bhu323

Brown, R. M., & Palmer, C. (2013). Auditory and motor imagery modulate learning in music performance. *Frontiers in Human Neuroscience*, *7*, 320. https://doi.org/10.3389/fnhum.2013.00320

Cirelli, L. K., Bosnyak, D., Manning, F. C., Spinelli, C., Marie, C., Fujioka, T., . . . Trainor, L. J. (2014). Beat-induced fluctuations in auditory cortical beta-band activity: Using EEG to measure age-related changes. *Frontiers in Psychology*, *5*, 742. https://doi.org/10.3389/fpsyg.2014.00742

Daly, I., Nicolaou, N., Williams, D., Hwang, F., Kirke, A., Miranda, E., & Nasuto, S. J. (2020). Neural and physiological data from participants listening to affective music. *Scientific Data*, *7*(1), 1–7. https://doi.org/10.1038/s41597-020-0507-6

Fukushima, K. (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, *36*(4), 193–202. https://doi.org/10.1007/bf00344251

Geuter, S., Lindquist, M. A., & Wager, T. D. (2017). Fundamentals of functional neuroimaging. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of Psychophysiology* (4th ed., pp. 41–73). Cambridge: Cambridge University Press.

González, M., Rojas, E., Bolaños, W., Segura, J. P., Murillo, L., Solano, A., González, E., Röhner, N., & Yu, L. (2019). Auditory imagery classification with a non-invasive brain computer interface. In *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER), San Francisco* (pp. 151–154). IEEE. doi: 10.1109/NER.2019.8716946

Herholz, S. C., Halpern, A. R., & Zatorre, R. J. (2012). Neuronal correlates of perception, imagery, and memory for familiar tunes. *Journal of Cognitive Neuroscience*, *24*(6), 1382–1397. https://doi.org/10.1162/jocn_a_00216

Huster, R. J., Debener, S., Eichele, T., & Herrmann, C. S. (2012). Methods for simultaneous EEG-fMRI: An introductory review. *Journal of Neuroscience*, *32*(18), 6053–6060. https://doi.org/10.1523/JNEUROSCI.0447-12.2012

Kingma, D. P., & Welling, M. (2013). Auto-encoding variational Bayes. arXiv preprint *arXiv preprint arXiv:1312.6114*.

Krishnan, A., & Plack, C. J. (2011). Neural encoding in the human brainstem relevant to the pitch of complex tones. *Hearing Research*, *275*(1–2), 110–119. https://doi.org/10.1016/j.heares.2010.12.008

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Losorelli, S., Nguyen, D. T., Dmochowski, J. P., & Kaneshiro, B. (2017). NMED-T: A tempo-focused dataset of cortical and behavioral responses to naturalistic music. In *Proceedings of the 18th International Society for Music Information Retrieval Conference (ISMIR)* (pp. 339–346). Suzhou: ISMIR. https://doi.org/10.5281/zenodo.1417917

Lotter, W., Kreiman, G., & Cox, D. (2016). Deep predictive coding networks for video prediction and unsupervised learning. *arXiv preprint arXiv:1605.08104*.

Moinnereau, M.-A., Brienne, T., Brodeur, S., Rouat, J., Whittingstall, K., & Plourde, E. (2018). Classification of auditory stimuli from EEG signals with a regulated recurrent neural network reservoir. *arXiv preprint arXiv:1804.10322*.

Nourski, K. V., Reale, R. A., Oya, H., Kawasaki, H., Kovach, C. K., Chen, H., Howard, M. A., & Brugge, J. F. (2009). Temporal envelope of time-compressed speech represented in the human auditory cortex. *Journal of Neuroscience*, *29*(49), 15564–15574.

Nozaradan, S., Peretz, I., & Mouraux, A. (2012). Selective neuronal entrainment to the beat and meter embedded in a musical rhythm. *Journal of Neuroscience*, *32*(49), 17572–17581. https://doi.org/10.1523/JNEUROSCI.3203-12.2012

Ofner, A., & Stober, S. (2018). Shared generative representation of auditory concepts and EEG to reconstruct perceived and imagined music. In *Proceedings of the 19th*

International Society for Music Information Retrieval Conference (ISMIR) (pp. 392–399). Paris: ISMIR. https://doi.org/10.5281/zenodo.1492433

Ofner, A., & Stober, S. (2019). Hybrid variational predictive coding as a bridge between human and artificial cognition. In *The 2018 Conference on Artificial Life (ALIFE)* (pp. 68–69). MIT Press Direct. https://doi.org/10.1162/isal_a_00142.

O'Sullivan, J. A., Power, A. J., Mesgarani, N., Rajaram, S., Foxe, J. J., Shinn-Cunningham, B. G., . . . Lalor, E. C. (2015). Attentional selection in a cocktail party environment can be decoded from single-trial EEG. *Cerebral Cortex*, *25*(7), 1697–1706. https://doi.org/10.1093/cercor/bht355

Pezzulo, G., Rigoli, F., & Friston, K. (2015). Active inference, homeostatic regulation and adaptive behavioural control. *Progress in Neurobiology*, *134*, 17–35. https://doi.org/10.1016/j.pneurobio.2015.09.001

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87. https://doi.org/10.1038/4580

Schaefer, R. S. (2014). Images of time: Temporal aspects of auditory and movement imagination. *Frontiers in Psychology*, *5*, 877. https://doi.org/10.3389/fpsyg.2014.00877

Schaefer, R. S., Desain, P., & Suppes, P. (2009). Structural decomposition of EEG signatures of melodic processing. *Biological Psychology*, *82*(3), 253–259. https://doi.org/10.1016/j.biopsycho.2009.08.004

Shahin, A., Roberts, L. E., Pantev, C., Trainor, L. J., & Ross, B. (2005). Modulation of p2 auditory-evoked responses by the spectral complexity of musical sounds. *Neuroreport*, *16*(16), 1781–1785. https://doi.org/10.1097/01.wnr.0000185017.29316.63

Shrestha, A., & Mahmood, A. (2019). Review of deep learning algorithms and architectures. *IEEE Access*, *7*, 53040–53065. https://doi.org/10.1109/access.2019.2912200

Stefanics, G., Heinzle, J., Horváth, A. A., & Stephan, K. E. (2018). Visual mismatch and predictive coding: A computational single-trial ERP study. *Journal of Neuroscience*, *38*(16), 4020–4030. https://doi.org/10.1523/JNEUROSCI.3365-17.2018

Sternin, A., Stober, S., Grahn, J., & Owen, A. (2015). Tempo estimation from the EEG signal during perception and imagination of music. In *1st International Workshop on Brain-Computer Music Interfacing/11th International Symposium on Computer Music Multidisciplinary Research (BCMI/CMMR'15), Plymouth*.

Stober, S. (2017). Toward studying music cognition with information retrieval techniques: Lessons learned from the openMIIR initiative. *Frontiers in Psychology*, *8*, 1255. https://doi.org/10.3389/fpsyg.2017.01255

Stober, S., Sternin, A., Owen, A. M., & Grahn, J. A. (2015). Towards music imagery information retrieval: Introducing the openMIIR dataset of EEG recordings from music perception and imagination. In *Proceedings of the 16th International Society for Music Information Retrieval Conference (ISMIR)* (pp. 763–769). Málaga: ISMIR. https://doi.org/10.5281/zenodo.1416270

Treder, M. S., Purwins, H., Miklody, D., Sturm, I., & Blankertz, B. (2014). Decoding auditory attention to instruments in polyphonic music using single-trial EEG classification. *Journal of Neural Engineering*, *11*(2), 026009. https://doi.org/10.1088/1741-2560/11/2/026009

Woodman, G. F. (2010). A brief introduction to the use of event-related potentials in studies of perception and attention. *Attention, Perception, & Psychophysics*, *72*(8), 2031–2046. https://doi.org/10.3758/BF03196680

10 Musical Imagery From a Cross-Cultural Methodological Perspective

George Athanasopoulos

Culture encompasses the social norms, behaviours, beliefs, knowledge, and universal values necessary for individuals to operate reasonably within a particular community (Little et al., 2014). It comprises imaginative and intellectual work and allows for the emergence of a unified mode of expression through integration of its people's various ways of life (e.g., consolidation of individual members' acquired habits and capabilities; Jenks, 2003). Culture also influences visual iconic and symbolic communicative characteristics, which, over time, and in combination with language, acquire meanings of their own. Understanding the manner in which these visual traits of iconic and symbolic communication function within communities in relation to other domains such as music requires careful examination of how the information included in the visual traits is communicated by members of said communities.

Musical imagery research in the domain of cross-cultural cognitive ethnomusicology requires a critical awareness of its larger situation within related disciplines, in particular musicology, psychology, and sociology. Without such awareness, researchers working in this field may be unable to anticipate how their initial epistemological positions, the formulation of their study, and the fieldwork setting may affect their results, as well as how said results will be interpreted by colleagues and the broader public. Lack of researchers' awareness, in turn, may bring about negative outcomes including exacerbation of scientific divisions or promotion of overtly sensational academic publications. Jacoby et al. (2020) highlight the need for researchers to integrate greater methodological rigor in cross-cultural research, especially when studying two or more communities that differ greatly in culture. Adopting effective and well-informed research strategies ensures theoretical robustness and leads towards successfully implementing accurate, credible, reliable, and dependable research (Lawson, 2014).

Starting Point

The starting challenge is to define *music*-related *mental imagery* in a way that is both comparable and meaningful across cultures. First, *music* presents many cross-cultural similarities (see Mithen et al., 2006; Trehub et al., 2015), as all documented cultures appear to have cultural practices involving some form of

DOI: 10.4324/9780429330070-13

organized sound. However, its manifestation may acquire different meanings in different cultures or even in the same culture across different times (Blacking, 1974). What is heard as music by an outside listener may not function as such within the society that produces the sound, as is usually the case with the Adhan (Muslim call to prayer) to non-familiar listeners—Muslim believers may appreciate oration techniques and styles, but under no circumstances would they consider the call to prayer to be "music." Secondly, *imagery* links visual descriptors to how humans make use of signs, indexes, and symbols metaphorically and literally, in relation to music (see Monelle, 2006) and may be further shaped by cultural constraints and affordances. Thirdly, the *mental* aspect (i.e., how specific parameters under observation may influence the mind) may be the most pernicious, as very early on in the preparation of a particular study, researchers are called to take a stance on which aspect of mental function they will focus on. Any pre-conceptions the researchers may have (e.g., whether cultural upbringing may affect participants' thought processes or whether presupposed universal underlying mechanisms may dictate mental processes) are also bound to have an effect. The majority of cross-cultural studies of music imagery will attempt either to trigger participants' episodic memory (i.e., by presenting music that participants may associate with a specific theme, context, or story; see also Jakubowski, this volume) or to investigate participants' implicit associations, as for example the association between core music parameters and spatial dimensions (i.e., pitch height and vertical height, or loudness and size). Studying implicit associations may be more useful, because episodic memory tends to differ among individuals *across* cultures as well as *within* a given culture (Wang et al., 2018). As such, it is impossible to locate participants across different cultures who are naïve to musical stimuli in a comparatively similar fashion, particularly when considering that musical structural knowledge can be implicitly acquired through mere exposure to a particular type of music (e.g., Tillmann et al., 2000; Wong et al., 2009).

Therefore, the objective of this chapter is to formulate a straightforward definition of mental musical imagery that can be employed while conducting research on cross-cultural musical imagery and the visualization of music. As the mental imagery of music is multimodal in nature (see Nanay, this volume), researchers are guided towards implementing primarily enactive visualizations in their research designs, associated with either aspects of music performance where visual imagery is linked to sensory/motor movement, action simulation, and/or movement execution (Keller, 2012), or, conversely, where visual imagery is represented using a visual/gesture depicting the overall musical surface as a two-/three-dimensional visual abstraction on a time frame. Consequently, mental musical imagery in a cross-cultural experimental setting may be conceived by participants as being a visual metaphor that is not frozen in time, but instead develops dynamically and in accordance with the music. The majority of studies dealing with the association between music and other domains (e.g., shape, size, emotion, colour, and action/movement) make use of language as an intermediate. These approaches often utilize tasks where (a) compatibility of music and a descriptor are assessed via self-report measures or (b) compatibility of music and specific (visual) concepts are assessed via timing response tasks (e.g., semantic priming). However, due to the very nature of language and its difficulty in being effectively translated across cultures, experimental designs relying on language-dependent metaphors (including those which are spatial, tactile, visual, and emotional) may yield a very wide range of responses and thus variable results (see Koelsch et al., 2004; Zhou et al., 2014, among others). In cross-cultural settings, researchers may rightly wish to avoid the linguistic element altogether, as translations may be inadequately comparable to their originals (e.g., different metaphors being used to describe the same sound in free response tasks; Dolscheid et al., 2013; Eitan & Timmers, 2010). Essentially, in order to perceive, and in turn, capture what participants from different cultural backgrounds experience as music imagery, linguistic metaphors have limited functionality in cross-cultural settings; therefore, by necessity, cross-cultural music research related to mental imagery could instead incorporate direct image representation without the use of language descriptors, presenting concepts via static iconic images or dynamic symbolic representations. In previous research, non-linguistic, iconic, and symbolic elements intended to visually represent sonic information in both free-drawn paradigms and forced-choice settings have been used cross-culturally with considerable success (see Antovic, 2009; Athanasopoulos et al., 2016; Athanasopoulos & Moran, 2013; Sadek, 1987; Walker, 1987), with results highlighting existing associations, or *conceptual blends*, drawn from participants' implicit memory. Conceptual blending between music and image concepts (Zbikowski, 2009) is consistent with a non-linear connectionist theory of cognitive processing, by which information across modalities is siphoned through multiple parallel processors and analysed in an integrated manner through mental networks connecting these processors. If it is possible to deconstruct musical referents and understand their component parts in terms of both their intra-musical (e.g., tempo, pitch height, and loudness) and their extra-musical parameters (e.g., age and gender of the performers, performance occasion, and performance location), it may accordingly be possible to envision these components via another sensory mode and prioritize those elements that are shared across cultures, ultimately resulting in comparable mental forms of music visualizations (Smith & Williams, 1997). Of course, people from different cultures may also produce *different* visualizations, based on their distinctive store of icono-symbolic referents. Similar pre-linguistic associations between music and images have also been suggested by other researchers (Evans & Treisman, 2010) introducing the question of whether cross-modal correspondences are innate or culturally acquired (Spence & Deroy, 2012). Spatial concepts have been assessed in the field of cross-cultural cognitive ethnomusicology numerous times, especially with regard to the association between pitch height y-axis position (for a detailed reference list, see Küssner et al., 2014), but also with regard to the association between loudness and verticality (Eitan et al., 2010) among others.

The body and its interaction with the physical environment have also been stipulated to play a key role during metaphorical representation of core musical parameters (see Eitan & Granot, 2006). Hatten (2017) holds the viewpoint that musical gesture is predominantly a biological affair, as the notion of giving shape to acoustic events in time is largely directed by the interplay between perceptual and motor systems, suggesting that non-linguistic conceptual metaphors connecting music and image may take on forms which closely resemble an organizational model of human knowledge based on multimodally blended domains (Godøy & Leman, 2010; Keil & Batterman, 1984; Malloch & Trevarthen, 2018).

At the same time, technological advancements of the 21st century have enhanced interconnectivity between people, making it challenging to judge and quantify exposure to non-native music and performance practices within cross-cultural research paradigms. It has also been argued that culture is not a reliable concept for the faithful description of permanently or objectively recognizable "domains" (Bashkow, 2004). Therefore, it is beyond reasonable limits to propose that data obtained via on-site fieldwork research are representative of universal human behaviour predating the overarching effect of cultural exchange. All cultures continuously interact with others around them, are influenced by them, and are constantly changing over time. Even if listening descriptors are avoided, how can we ask participants to develop music-related images and communicate information about them effectively? Moreover, how can we encourage them to include aspects of their culture while doing so? As noted by Jacoby et al. (2020), it is crucial to acknowledge that research participants, in today's digital world, may have listening experiences which cross their cultural and geographical borders. Since this is now true for even the most remote places around the globe, it is essential for researchers to carefully identify their assumptions before project proposal formation, such that it can be determined whether embarking on a fieldwork trip to collect data from a particular group is necessary or whether similar data can be collected via online crowdsourcing or from migrant communities within the researchers' vicinity.

Moving Forward: Ethics, and Data From the Field

Ethics

For prospective researchers in the field, it is instrumental to consider pursuing a multidisciplinary approach when investigating cross-modal musical concepts. This should be done not only in the interpretation of the results (e.g., when analysing musical shapes and two-dimensional representations put forth by participants) but also during the preparation stage starting with what research questions are being asked, in which manner, order, and whether they make sense in terms of the participants' cultural context(s). Before any research questions are posed, ethical considerations should be taken into account. Ethics cannot be separated from methodology; they inform each other. Research in cross-cultural music-related mental imagery raises ethical concerns that are no less important than those concerns raised by other psychological research involving human participants. Researchers must recognize that they are often in a position of power, and accordingly, should minimize any negative effect that this power dynamic may have on participants, for instance, by emphasizing *reciprocity* in the researcher–participant relationship and ensuring that participants be properly compensated for their time. Academic institutions, whose reviewing teams ought to be made up of scholars of different backgrounds and work closely with the institution's international advisory office, should scrutinize research projects proposed by staff members and students for methodological rigour and ethical approach to data collection. This is critical for ensuring not only that construction of the method section and acquisition of data are maximally respectful of persons from other cultures, but also that the research question itself carefully addresses the cultural impact that the research project may have on its participants. Researchers should be wary of introducing novel ideas through their research projects (i.e., introducing ideas to the culture they are researching that are not already present in that culture), as this may have adverse effects on any participating group—for example, introducing the idea of visually representing music to a non-literate population might have unforeseen consequences, such as contributing to an oral tradition being phased out or disrupting existing modes of music learning.

Data From the Field

The data from the field presented here are from cross-cultural experiments that I conducted with a number of collaborators, involving participants in the United Kingdom, Greece, Japan, Papua New Guinea (Athanasopoulos et al., 2016; Athanasopoulos & Antović, 2018; Athanasopoulos & Kitsios, 2015; Athanasopoulos & Moran, 2013), and most recently from Pakistan, each experiment using comparably similar stimuli, participants' prompts, and self-report response items (i.e., free-drawn and forced-choice representations). The results that I have presented over the years suggest that humans' thought processes may indeed share innate associations between sonic events and pre-existing imagery concepts, and such associations are likely to operate in accordance with their own cultural practices. Though the vertical representation of pitch crucially underpins this cross-modal association (see Cox, 2016), further reasons for its existence are multiple and varied (for a detailed analysis, see Eitan, 2017). Consequently, it may be inferred that the symbolic Cartesian representation most often depicted by participants in the United Kingdom, Greece, and Japan in the aforementioned studies arises due to embodied, enactive, and cultural parameters. Conversely, other participants originate from countries where depicting musical/lexical information in time is not the norm, or whose representational systems are intrinsically different from modes of depiction akin to Western Standard Notation (WSN), as for example, Noh musicians from Japan. Nohkan, tsuzumi, and vocal notations from Noh theatre all run vertically from right to left, and are extremely instrument-specific. Such participants who do not have notational systems as part of their music traditions may be less internally consistent; nevertheless, they are fully capable of producing and incorporating musical imagery despite the lack of pre-existing associations with a versatile written representational mode.

As evinced by participant data from tasks that link music to a representational form and bypass linguistic metaphors (i.e., free-drawing paradigms which *do* *not* involve pre-existing methods of musical representation such as music notation), music-related mental imagery appears to be linked to the cognitive aspect of metaphorical gestures involving path and movement in spatial locatives as a description of sonic events, or coded as a prescriptive "performance" guide, similar to Seeger's classification (1958). Although motor output is usually inhibited due to experimental constraints, it can be easily disinhibited when movement is envisioned internally (i.e., as a top-down thought process), resulting in symbolic (arbitrary connections between an image and its meaning) horizontal left-to-right representations. Still, necessary measures must be taken in order to avoid bias at all stages of the experiment (e.g., sampling bias; for a full list, see Pannucci & Wilkins, 2010). Recruiting adult participants trained in or familiar with music from a wide and varied selection of cultures around the world somewhat mitigates but does not entirely eliminate this sampling bias; this is due to the fact that even in multiple culture studies, the samples chosen may not be appropriate for answering the questions posed during research. Studying musical imagery under a cross-cultural methodological perspective requires acknowledgement of the uniquely complex concerns pertaining to a given participant pool and the consequent adoption of equally unique effective and ethically sound methods. This research approach requires thorough examination of cultural issues stemming from ethnomusicology, music, and imagery, and other related fields that may inform such studies. When using a cross-cultural methodology, the unreliability (or worse: non-existence) of a conceptual common ground may pose a major challenge for both researchers and participants. This issue may result in adverse miscommunications while pursuing research and/or communicating findings, or even the production of research with little to no relevance for the cultures from which a given participant originates from.

Discussion

One of the key issues when interpreting results in cross-cultural psychology research is that researchers often misconstrue or neglect the comprehensive investigation of culture, how it relates to psychology, and its interaction with biological processes; this may be due to the inadequate fit of their research question's theoretical grounding to the cultural considerations of the population that they are investigating (Ratner, 2003). Such neglected considerations, which further include failure to thoroughly examine the interaction between culture and biological factors, may thwart one's ability to produce comprehensive and astute observations about the data. Researchers engaging in cross-cultural studies should acknowledge that there is a very real possibility of misconceiving some cultural elements; admitting and accounting for possible weaknesses are the only ways to facilitate production of quality research, and to proactively mitigate errors (Halder et al., 2016; Stevens, 2012; Trehub, 2015). Furthermore, interdisciplinary collaboration among researchers is likely to improve such cross-cultural projects; incorporating anthropologists and ethnomusicologists into cognitive musicologists' research teams may absolve methodological flaws as well as misinterpretations of data.

Looking at the fieldwork data obtained from the aforementioned series of experiments, it is worthwhile to discuss which cultural parameters affect participants' two-dimensional responses in musical imagery experiments. It is critical that the experimental procedure incorporates self-report measures (see also Hubbard, this volume) that allow participants to accurately convey their own internal visual representation incited by a given musical stimulus. In formulating such measures, some general guidelines emerge. First, any type of formal school experience—involving any method of transmission, and in any cultural setting—is important in that it enables participants to reciprocate and react to external instructions, and to become energetically involved in any task involving research participation. In this context, schooling should be understood as any method of transmitting knowledge from a tutor to a single or to a group of individuals. Secondly, music training on any musical instrument appears to have a positive effect on the comprehension of research tasks (Haning, 2016), and on the consistency of participants' responses, particularly when the response involves associating music with images (Küssner et al., 2014). Thirdly, while literacy does have an effect on participant responses, particularly for organizational aspects of the visual representations (e.g., whether or not they are along an axis; see Athanasopoulos et al., 2011, 2016), representations of participants' responses may also vary in other respects. For instance, non-literate participants tend to focus on representing timbre, while at the same time attempting to depict instructional "scores," likely similar to the manner by which they instruct their peers to re-create the sonic stimuli. This is particularly true for representations of attack rate among non-literate Papua New Guineans (Athanasopoulos et al., 2016). Fourthly, while researchers often prioritize the construction of musical stimuli so that participants can classify them as (non-genre-specific) music which may or may not be from their own culture and using a timbre which is familiar to them, it is more likely that problems will arise due to physical aspects of the experimental setup—for example, if a participant has never worn headphones up to that point in their adult life, or has never seen a Galvanic Skin Response or a Heart-Rate Variation sensor, the stress of the experience alone is likely to impact their responses. Fifthly, participants' unfamiliarity with the data collection method (e.g., asking participants not familiar with a computer mouse-pad to deliver timed responses) may confound the obtained results. Sixthly, the application of ambiguous scale items, particularly in forced-choice tasks where participants need to match sonic information with imagery or word descriptors, may also lead to faulty responses. The aforementioned issues, especially when compounded, may bring about later difficulties when attempting to accurately conduct cross-cultural fieldwork and draw meaningful conclusions from the data. Determining potential challenges and designing measures of addressing them at an earlier point is necessary in order to avoid such pitfalls (Ratner, 2003). An additional consideration that should be made, as ascertained from collected participant data, is that while coded musical information (i.e., symbolic, left-to-right horizontal representation) appears to be more common among musically literate participants, participants' representational techniques should not be seen as points across a linear "progression." The conceptual blend between visual representations and music does not appear to be a gradual evolutionary progress from one stage (iconic) to another (symbolic). There are no "primitives," there is no infancy, adolescence, or maturity. On the contrary, what appears to be the case is that different participant groups seem to be employing different combinations of (culturally determined) metaphorical or literal concepts when listening to music, and that cross-cultural commonalities, such as adoption of a timeline when depicting musical information, may appear (Godøy & Leman, 2010; Trehub et al., 2015). Therefore, as outlined here, and further supported by empirical evidence presented in cross-cultural research in music as two-dimensional representation (i.e., Athanasopoulos & Moran, 2013; Athanasopoulos et al., 2016), there is no objective quality to the perception of musical imagery; instead, each participant group is likely utilizing approaches which maximally facilitate communication between themselves and their peers.

Conclusion

It is apparent that a cross-cultural examination of musical imagery includes numerous challenges. The definition of musical imagery is linked to numerous cultural elements and variable parameters, including but not limited to metaphorical concepts, social norms, behaviours, and values that inform how people produce mental images of sound. An interdisciplinary approach in conducting cross-cultural research is necessary, and ideally includes incorporation of concepts, approaches, and methods from anthropology, psychology, sociology, and ethnomusicology. Combining different methodological approaches, as well as carefully planning experiments, may improve a given researcher's methodological rigour, whereas acknowledging the complexity of such research and the challenges faced by cross-cultural researchers will help to formulate a widespread methodology that allows more comprehensive, in-depth, and meaningful research to be carried out.

Research across a wide range of cultural contexts is vital for establishing the relevance of musical imagery to a broader population. As demonstrated here, the challenges of conducting cross-cultural research in music and imagery are likely to persist and require the use of effective approaches that incorporate a broader account of cultural sensitivities so as to include their potential influence on the research topic at hand and enhance the credibility of the research's results. It is important to note that all cross-cultural research in the humanities and social sciences is quite challenging. The present chapter has demonstrated that the potentials and constraints of empirical methodological applications are numerous when conducting cross-cultural research in music and visual imagery in terms of both its structure and its implementation, and even if vigorous, thorough, and culturally appropriate pre-assessment tests are included. At the same time, research in this domain is likely to be extremely rewarding and to produce results that can potentially enhance our understanding of human thought processing and cross-modal perception in music-related mental imagery. Still, for researchers ambitious enough to engage in such projects, proactive preparation beforehand as outlined earlier is absolutely necessary for any chance of success.

References

Antovic, M. (2009). Musical metaphors in Serbian and Romani children: An empirical study. *Metaphor and Symbol*, *24*(3), 184–202. https://doi.org/10.1080/10926480903028136

Athanasopoulos, G., & Antović, M. (2018). Conceptual integration of sound and image: A model of perceptual modalities. *Musicae Scientiae*, *22*(1), 72–87. https://doi.org/10.1177/1029864917713244

Athanasopoulos, G., & Kitsios, G. (2015). Sonic perception translated into image: Greek classically-trained/traditional musicians attempt to portray classical/traditional music as graphic representations. In *Proceedings of the 2015 musicult—locality and universality* (pp. 95–106). Istanbul: Istanbul Technical University.

Athanasopoulos, G., & Moran, N. (2013). Cross-cultural representations of musical shape. *Empirical Musicology Review*, *8*(3–4), 185–199. https://doi.org/10.18061/emr.v8i3-4.3940

Athanasopoulos, G., Moran, N., & Frith, S. (2011, August 11). *Literacy makes a difference: A cross-cultural study on the graphic representation of music by communities in the United Kingdom, Japan and Papua New Guinea* [Conference Presentation]. Biennial Meeting of the Society for Music Perception and Cognition (SMPC). Rochester, NY.

Athanasopoulos, G., Tan, S. L., & Moran, N. (2016). Influence of literacy on representation of time in musical stimuli: An exploratory cross-cultural study in the UK, Japan, and Papua New Guinea. *Psychology of Music*, *44*(5), 1126–1144. https://doi.org/10.1177/0305735615613427

Bashkow, I. (2004). A neo-Boasian conception of cultural boundaries. *American Anthropologist*, *106*(3), 443–458. https://doi.org/10.1525/aa.2004.106.3.443

Blacking, J. (1974). *How musical is man?* Washington: University of Washington Press.

Cox, A. (2016). *Music and embodied cognition: Listening, moving, feeling, and thinking*. Bloomington: Indiana University Press.

Dolscheid, S., Shayan, S., Majid, A., & Casasanto, D. (2013). The thickness of musical pitch: Psychophysical evidence for linguistic relativity. *Psychological Science*, *24*(5), 613–621. https://doi.org/10.1177/0956797612457374

Eitan, Z. (2017). Musical connections: Cross-modal correspondences. In R. Ashley & R. Timmers (Eds.), *The Routledge companion to music cognition* (pp. 213–224). New York: Routledge.

Eitan, Z., & Granot, R. Y. (2006). How music moves: Musical parameters and listeners images of motion. *Music Perception*, *23*(3), 221–248. https://doi.org/10.1525/mp.2006.23.3.221

Eitan, Z., Schupak, A., & Marks, L. E. (2010). Louder is higher: Cross-modal interaction of loudness change and vertical motion in speeded classification. In K. Miyazaki, Y. Hiraga, M. Adachi, Y. Nakajima, & M. Tsuzaki (Eds.), *Proceedings of the 10th International Conference on Music Perception and Cognition (ICMP10)* (pp. 67–76). Sapporo: ICMPC10.

Eitan, Z., & Timmers, R. (2010). Beethoven's last piano sonata and those who follow crocodiles: Cross-domain mappings of auditory pitch in a musical context. *Cognition*, *114*(3), 405–422. https://doi.org/10.1016/j.cognition.2009.10.013

Evans, K. K., & Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features. *Journal of Vision*, *10*(1), 1–12. doi:10.1167/10.1.6

Godøy, R. I., & Leman, M. (2010). *Musical gestures: Sound, movement, and meaning*. London: Routledge.

Halder, M., Binder, J., Stiller, J., & Gregson, M. (2016). An overview of the challenges faced during cross-cultural research. *Enquire*, *8*, 1–18.

Haning, M. (2016). The association between music training, background music, and adult reading comprehension. *Contributions to Music Education*, *41*, 131–143.

Hatten, R. S. (2017). A theory of musical gesture and its application to Beethoven and Schubert. In A. Gritten & E. King (Eds.), *Music and gesture* (pp. 1–18). Abingdon: Routledge. https://doi.org/10.4324/9781315091006-2

Jacoby, N., Margulis, E. H., Clayton, M., Honing, H., Iversen, J., Klein, T., . . . Wald-Fuhrmann, M. (2020). Cross-cultural work in music cognition: Challenges, insights, and recommendations. *Music Perception*, *37*(3), 185–195. doi:10.1525/mp.2020.37.3.185

Jenks, C. (2003). *Culture: Critical concepts in sociology*. London: Taylor & Francis.

Keil, F. C., & Batterman, N. (1984). A characteristic-to-defining shift in the development of word meaning. *Journal of Verbal Learning and Verbal Behavior*, *23*(2), 221–236. https://doi.org/10.1016/s0022-5371(84)90148-8

Keller, P. E. (2012). Mental imagery in music performance: underlying mechanisms and potential benefits. *Annals of the New York Academy of Sciences*, *1252*(1), 206–213. https://doi.org/10.1111/j.1749-6632.2011.06439.x

Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., & Friederici, A. D. (2004). Music, language and meaning: Brain signatures of semantic processing. *Nature Neuroscience*, *7*(3), 302–307. https://doi.org/10.1038/nn1197

Küssner, M. B., Tidhar, D., Prior, H. M., & Leech-Wilkinson, D. (2014). Musicians are more consistent: Gestural cross-modal mappings of pitch, loudness and tempo in real-time. *Frontiers in Psychology*, *5*, 789. https://doi.org/10.3389/fpsyg.2014.00789

Lawson, F. R. (2014). Is music an adaptation or a technology? Ethnomusicological perspectives from the analysis of Chinese Shuochang. *Ethnomusicology Forum*, *23*(1), 3–26. https://doi.org/10.1080/17411912.2013.875786

Little, W., Vyain, S., Scaramuzzo, G., Cody-Rydzewski, S., Griffiths, H., Strayer, E., & Keirns, N. (2014). *Introduction to sociology—1st Canadian edition*. Victoria: BC Campus.

Malloch, S., & Trevarthen, C. (2018). The human nature of music. *Frontiers in Psychology*, *9*(1680), 1–21. https://doi.org/10.3389/fpsyg.2018.01680

Mithen, S., Morley, I., Wray, A., Tallerman, M., & Gamble, C. (2006). The singing Neanderthals: The origins of music, language, mind and body, by Steven Mithen. London: Weidenfeld & Nicholson, 2005; ix+374 pp. [Review]. *Cambridge Archaeological Journal*, *16*(1), 97–112. https://doi.org/10.1017/S0959774306000060

Monelle, R. (2006). *The musical topic: Hunt, military and pastoral*. Bloomington: Indiana University Press.

Pannucci, C. J., & Wilkins, E. G. (2010). Identifying and avoiding bias in research. *Plastic and Reconstructive Surgery*, *126*(2), 619. https://doi.org/10.1097/prs.0b013e3181de24bc

Ratner, C. (2003). Theoretical and methodological problems in cross-cultural psychology. *Journal for the Theory of Social Behavior*, *33*(1), 67–94. https://doi.org/10.1111/1468-5914.00206

Sadek, A. A. M. (1987). Visualization of musical concepts. *Bulletin of the Council for Research in Music Education*, 149–154.

Seeger, C. (1958). Prescriptive and descriptive music-writing. *The Musical Quarterly*, *44*(2), 184–195. https://doi.org/10.1093/mq/xliv.2.184

Smith, S. M., & Williams, G. N. (1997, October). A visualization of music. In *Proceedings: Visualization '97 (Cat. No. 97CB36155)* (pp. 499–503). Phoenix, AZ: IEEE.

Spence, C., & Deroy, O. (2012). Crossmodal correspondences: innate or learned? *Iperception*, *3*(5), 316–318. doi:10.1068/i0526ic

Stevens, C. J. (2012). Music perception and cognition: A review of recent cross-cultural research. *Topics in Cognitive Science*, *4*(4), 653–657. https://doi.org/10.1111/j.1756-8765.2012.01215.x

Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review*, *107*(4), 885–913. https://doi.org/10.1037/0033-295x.107.4.885

Trehub, S. E. (2015). Cross-cultural convergence of musical features. *Proceedings of the National Academy of Sciences*, *112*(29), 8809–8810. https://doi.org/10.1073/pnas.1510724112

Trehub, S. E., Becker, J., & Morley, I. (2015). Cross-cultural perspectives on music and musicality. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *370*(1664), 20140096. https://doi.org/10.1098/rstb.2014.0096

Walker, R. (1987). The effects of culture, environment, age, and musical training on choices of visual metaphors for sound. *Perception & Psychophysics*, *42*(5), 491–502. https://doi.org/10.3758/bf03209757

Wang, Q., Hou, Y., Koh, J. B. K., Song, Q., & Yang, Y. (2018). Culturally motivated remembering: The moderating role of culture for the relation of episodic memory to well-being. *Clinical Psychological Science*, *6*(6), 860–871. https://doi.org/10.1177/2167702618784012

Wong, P. C., Roy, A. K., & Margulis, E. H. (2009). Bimusicalism: The implicit dual enculturation of cognitive and affective systems. *Music Perception*, *27*(2), 81–88. https://doi.org/10.1525/mp.2009.27.2.81

Zbikowski, L. (2009). Music, language, and multimodal metaphor. In C. Forceville & E. Urios-Aparisi (Eds.), *Multimodal metaphor* (pp. 359–382). Berlin, NY: De Gruyter Mouton. https://doi.org/10.1515/9783110215366.6.359

Zhou, L., Jiang, C., Delogu, F., & Yang, Y. (2014). Spatial conceptual associations between music and pictures as revealed by N400 effect. *Psychophysiology*, *51*(6), 520–528. https://doi.org/10.1111/psyp.12195

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Part III

Mental Imagery and Related States of Consciousness

11 Mental Imagery in Music-Evoked Autobiographical Memories

Kelly Jakubowski

The experience of being mentally transported back to a previous event or time period from one's life when listening to music is a familiar one to many people. Such music-evoked autobiographical memories (MEAMs) are typically vivid and emotionally positive (Belfi et al., 2016; Janata et al., 2007), and tend to occur around once per day (Jakubowski & Ghosh, 2019). MEAMs often come to mind involuntarily, with no intentional effort made to retrieve a memory (El Haj et al., 2012a; Jakubowski & Ghosh, 2019), and can vary in their level of specificity (Janata et al., 2007; Kristen-Antonow, 2019)—from semantic information or facts about one's life, to generic memories of entire life periods or recurring events, to specific, episodic details about a single event (see Conway & Pleydell-Pearce, 2000, for one influential model of the levels of organization of autobiographical knowledge). The main autobiographical memories of interest for the present chapter are those that contain at least some episodic content, meaning that these memories are characterized by a sense of recollection and "reliving" of past experiences, which is made possible via mental imagery. In this chapter, I will first provide a brief overview of research on mental imagery in autobiographical memory in general, followed by a review of relevant findings on mental imagery in MEAMs, and suggestions for future research.

Mental Imagery in Autobiographical Memories

In research on autobiographical memory more broadly, the role of mental imagery in recollective experiences has been explored in some detail. *Mental imagery* will be defined in this chapter as a perception-like experience that can occur even in the absence of perceptual input. Experiences of mental imagery inevitably rely on memory processes (e.g., "imagine the face/voice of your mother"), but can also recombine previously perceived stimuli in novel or creative ways (e.g., "imagine your mother speaking in your father's voice"). This possibility for recombination in mental imagery bears many parallels to discussions of the reconstructive nature of autobiographical memory (i.e., memories are not an exact record of the past, but can change over time and are influenced by various biases and errors; Hyman & Loftus, 1998). Mental imagery can exist in a single modality (visual, auditory, haptic, olfactory, etc.), but it can also be multimodal (Nanay, this volume).

DOI: 10.4324/9780429330070-15

Studies of mental imagery in autobiographical memories to date have focused predominantly on the visual modality (Taruffi & Küssner, this volume), with some taxonomies of autobiographical memory including visual imagery as one of the defining features by which episodic autobiographical memories can be distinguished from semantic autobiographical facts (e.g., being able to mentally visualize your sixth birthday party at a particular playground, vs. simply knowing that your sixth birthday party happened at that playground; Brewer, 1986).

Indeed, subsequent research has indicated that visual imagery is a strong predictor of ratings of the strength of recollection in autobiographical memories, although it is clear that these memories can also be accompanied by mental imagery across several different sensory modalities (e.g., auditory and olfactory imagery; Rubin, 2005). Rubin et al. (2003a) studied autobiographical memories recalled by undergraduate students in response to word cues, such as "tree" and "doctor" (following Crovitz & Schiffman, 1974). The researchers analysed the extent to which participant ratings of reliving, remembering an event (rather than just knowing it happened), and mental time travel were predicted by component properties of the memory. Ratings of visual imagery and, to a lesser extent, auditory imagery were both significant predictors of these measures of the recollective nature of the memories. In addition, memories that were reported as highly relived were almost always (98% of the time) accompanied by strong visual images. These results provide compelling evidence that mental imagery, particularly in the visual domain, may be a necessary component for the recollection of autobiographical events. This account has been supported by a series of experiments demonstrating that an absence of visual input at encoding (e.g., by blindfolding participants or presenting video recordings in audio-only versions) significantly reduced ratings of recollection of events at retrieval (Rubin et al., 2003b). Individual differences in visual imagery abilities have also been found to be predictive of the level of sensory-perceptual detail and specificity of autobiographical memories (Aydin, 2018), as well as the way in which event and spatial details are remembered (Sheldon et al., 2017). Finally, damage to brain regions necessary for visual memory has been found to be accompanied by profound impairments in autobiographical recall, whereas linguistic and auditory impairments are much less associated with autobiographical memory deficits (Greenberg & Rubin, 2003).

Subsequent work has highlighted that episodic memory recall activates highly similar brain networks to several other cognitive activities, including thinking about the future, imagining fictitious experiences, navigation, and mind-wandering (Hassabis & Maguire, 2007; Konishi, this volume).[1] Hassabis and Maguire (2007) have thus proposed "scene construction" as a crucial component process underlying all these cognitive functions. Scene construction is defined as "the process of mentally generating and maintaining a complex and coherent scene or event," involving retrieval and integration of multimodal sensory information into a spatial context that may subsequently be manipulated or transformed (Hassabis & Maguire, 2007, p. 299). Scene construction provides a quite parsimonious explanation for the striking similarities between a seemingly diverse set of cognitive functions, all of which implicate the generation and manipulation of multimodal imagery. Thus, it appears that the component processes underlying autobiographical memories may overlap substantially with other cognitive activities outlined in this book, including mind-wandering, imagination, and mental simulation, with autobiographical memories being differentiated primarily on the basis of their evocation of a sense of mental time travel (a key component of "autonoetic consciousness"; Tulving, 1985), past-focused temporal orientation, and connection to the self.

Mental Imagery in Music-Evoked Autobiographical Memories

More recently, researchers have become interested in the specific experience of autobiographical memories cued by music, in particular as music appears to be an effective means for spontaneously eliciting positive and significant lifetime memories (El Haj et al., 2012a; Janata et al., 2007). MEAMs are a topic of great theoretical interest in terms of understanding the links between music, emotions, and identity, and practical interest in terms of exploring the potential for using music to elicit memories in people with dementia and other memory impairments. Although there is still much to be discovered regarding the mechanisms that underlie the coupling of music to autobiographical memories, several studies have revealed initial insights regarding mental imagery in MEAMs, and in particular, how MEAMs might differ from other autobiographical memories in terms of imagery experiences.

The Content of Mental Imagery in Music-Evoked Autobiographical Memories

In a seminal study, Janata et al. (2007) elicited MEAMs by playing a wide range of chart-topping pop songs to undergraduate students. Their work presents a detailed characterization of the MEAM experience, including some first indications of their imaginal content. On average, over 40% of the songs that were played elicited memories of people—most typically friends and significant others—and for 83% of participants at least one song elicited a memory of a person. Fewer songs evoked memories of places (less than 20% of songs on average). Written descriptions of the memory contents were analysed using Linguistic Inquiry and Word Count (LIWC; Pennebaker et al., 2003), a software that performs automated categorization of words into themes using large dictionaries of conceptually related words. The most highly represented word categories were the "Social" and "Leisure" categories, and the most commonly used words were "school," "friend(s)," "danc(e/es/ing)," and "car/driving."

Jakubowski and Ghosh (2019) subsequently investigated the content and features of naturally occurring MEAMs, using a week-long diary method. MEAM descriptions were classified using LIWC and were again characterized by a relatively high degree of "Social" and "Leisure" category words, and the most commonly used words bore strong similarities to those found by Janata et al. (2007).

MEAMs also exhibited a relatively high percentage of words in the "Hear" LIWC category (a subcategory of "Perceptual processes"); this seemed to be mainly due to a high usage of words such as "song," "listen," and "radio," in which memory descriptions often referred to previous instances of listening to the same song. However, on average only around 1% of words from each MEAM description fell into the "See" category from LIWC. It is yet to be determined whether this result indicates a relatively low degree of visual imagery within MEAMs, or whether the LIWC dictionary simply does not capture linguistic descriptions in a way that can be meaningfully interpreted as evidence of visual imagery. In addition, although these studies provide some initial overview and agreement in terms of the typical content of MEAMs, further research is needed to test whether such memories actually evoke mental imagery (e.g., whether music brings to mind an image of your schoolfriend Jill, or simply reminds you that you had a schoolfriend named Jill) and, if so, what modality such mental imagery might take (e.g., do you simply remember Jill's face, or the sound of her voice and the smell of her perfume?).

Comparing Mental Imagery in Music-Evoked Autobiographical Memories to Other Autobiographical Memories

Several studies have compared the phenomenological characteristics or content of MEAMs to autobiographical memories evoked via other cues. Belfi et al. (2016) compared MEAMs cued by pop songs to autobiographical memories evoked by photographs of famous faces. Memory descriptions were coded using the procedures of Levine et al. (2002), and it was found that MEAMs contained a greater proportion of episodic and perceptual details than memories evoked via famous faces, indicating that MEAMs were characterized by more vivid and detailed reliving of the remembered events. A subsequent study replicated these findings, and also revealed that MEAMs were significantly less episodically rich in patients with damage to the medial prefrontal cortex (mPFC) in comparison to healthy controls (Belfi et al., 2018). The patients did not differ from the controls in terms of the episodic richness of memories evoked by famous faces, suggesting a specific role for the mPFC in integrating musical cues with episodic memory details.

Zator and Katz (2017) compared memories cued by popular music to memories evoked via two types of word cues, which referred to either a life period or a specific event. They used LIWC to analyse the written memory descriptions and found that the event-specific word cues elicited significantly more words from the "See" LIWC category, while musical cues elicited a significantly higher proportion of words in the "Hear" category, indicating potential differences in the modality of mental imagery implicated in these two memory types. MEAM descriptions also contained significantly more motor and spatial terms (indexed by the "Relativity," "Motion," and "Space" LIWC categories) than the other two memory types, suggesting that music cued memories that were more embodied and perhaps accompanied by increased visuospatial and motor imagery.

Jakubowski et al. (2021) compared MEAMs to autobiographical memories evoked by watching television (TV) in an online survey of a representative sample of UK adults. MEAMs were rated significantly higher in vividness and reliving and contained more perceptual and social details (as coded using LIWC) than TV-evoked memories; these differences were also consistent across three age groups (young, middle-aged, and older adults). These results occurred despite the fact that the MEAMs and TV-evoked memories did not differ in terms of how recently they had been recalled, and the music and TV programmes also did not differ in terms of how much they were liked by the participants. However, the music was rated as significantly more familiar than the TV programmes, suggesting that the association between a piece of music and its corresponding autobiographical memory may have been more frequently rehearsed, leading to increases in the amount of episodic detail recalled.

These studies provide some initial evidence on how the imaginal contents of MEAMs differ from other autobiographical memories. The language used in memory descriptions has indicated that MEAMs differ from other memories in terms of their increased embodied nature and social content (Jakubowski et al., 2021; Zator & Katz, 2017). In addition, MEAM descriptions have been found to comprise a greater proportion of perceptual details than memories cued by photographs of famous faces and TV programmes (Belfi et al., 2016; Jakubowski et al., 2021).

One question that remains is *why* music is able to bring back more episodically detailed memories than other common memory cues. As mentioned earlier, pieces of music are often listened to over and over for many years, which may strengthen the association between the music and an autobiographical memory. Although the cue overload principle (Berntsen, 2009) suggests that hearing the same piece of music in the context of many different life events can also *blur* the association between the musical cue and a particular memory, this may be partially counteracted by the fact that some instances of music listening are coupled with life events that are highly emotional and can be seen as milestones or "turning points" (i.e., music is often heard at weddings, funerals, and initiations). These particularly salient and important memories are less likely to be susceptible to cue overload, especially if future instances of exposure to the associated piece of music occur in relatively mundane, non-emotional contexts. The relationship between the emotional properties of a musical retrieval cue and the emotional reaction to the associated memory also merits further exploration, in terms of how emotional features of the cue itself might enhance recall of certain episodic details (see initial exploration of these ideas in Schulkind & Woldorf, 2005, and Sheldon & Donahue, 2017).

In addition, several studies have highlighted the involuntary nature of MEAMs. Involuntary memories are retrieved in a direct and spontaneous manner, rather than via an intentional and effortful search process. Self-selected music appears to be particularly effective for eliciting involuntary autobiographical memories (El Haj et al., 2012a); this parallels results on other types of autobiographical memories, in which personal/self-selected cues elicited more directly retrieved memories than experimenter-selected cues (e.g., Uzer & Brown, 2017). Some studies (e.g., Cuddy et al., 2017) have even defined MEAMs as a specific type of involuntary memory, although other research indicates that the majority of MEAMs are involuntary, but voluntary MEAMs can also occur in everyday life (Jakubowski & Ghosh, 2019).[2] Research on involuntary autobiographical memories has revealed that such memories are generally characterized by greater reliving, including heightened emotional impact, than voluntarily retrieved autobiographical memories (Berntsen & Hall, 2004). This work also indicates that involuntary retrieval is more effective for accessing memories of specific episodes, whereas voluntary retrieval tends to favour more generic memories. Interestingly, however, Jakubowski et al. (2021) did not find a significant difference in ratings of the involuntary nature of MEAMs versus TV-evoked memories, despite increased ratings of reliving and vividness in MEAMs. Thus, it seems that involuntary retrieval may be only part of the explanation for differences in recall of episodic detail.

Conclusions and Future Research

Although research on autobiographical memory has broadly examined the role of mental imagery (especially visual imagery) in recollective experiences, the imaginal processes implicated in MEAMs have only recently been explored. Initial empirical findings have indicated that MEAMs of healthy adults may be relived in greater episodic detail than autobiographical memories evoked via other pop cultural cues (Belfi et al., 2016, 2018; Jakubowski et al., 2021). However, the underlying mechanisms that might explain such differences (e.g., encoding conditions, retrieval method, rehearsal frequency, and emotional factors) are still poorly understood.

Several studies have also provided indications of the content of mental imagery in MEAMs. MEAMs often comprise social themes, including memories of people such as friends and significant others (Jakubowski & Ghosh, 2019; Jakubowski et al., 2021; Janata et al., 2007). This finding speaks to the prominent role music can play in developing and maintaining social bonds, including group formation on the basis of shared personal preferences and values (e.g., Rentfrow & Gosling, 2006). Descriptions of MEAMs were also characterized by a relatively high proportion of motor and spatial language (Zator & Katz, 2017), which relates to the embodied nature of music listening. This fits with theories emphasizing the close coupling of perception and action systems during music listening (Maes et al., 2014), including the common urge to move along with music (Janata et al., 2012). It may be that this drive to action could actually serve as an additional cue for remembering a previous experience involving moving to music; indeed, words such as “dance” and “sing” were particularly prevalent in the studies of both Janata et al. (2007) and Jakubowski and Ghosh (2019). One potential avenue for subsequent research could be to explore whether music that is high in groove (“the urge to move in response to music, combined with the positive affect associated with the coupling of sensory and motor processes”; Janata et al., 2012, p. 54) elicits more embodied and vivid autobiographical memories than low-groove music.

The studies reviewed earlier have typically drawn inferences about the imaginal contents of MEAMs based on the language used by participants to describe their memories; they therefore provide somewhat indirect evidence of the underlying mental imagery processes. Future research should more explicitly ask participants to report on their imagery experiences during MEAMs (Hubbard, this volume), including details on the modality of such imagery. More covert measures may also provide new insights (Eerola, this volume), for instance by examining the activation of visual imagery-related brain areas during MEAMs or by attempting to suppress visual imagery during MEAMs to test the effects of such a manipulation on the quality of the recollective experience (Belfi, this volume). Studies of mental imagery in modalities other than the visual domain are relatively rare in autobiographical memory research, but will also be important for providing a more complete understanding of these everyday mental experiences.

One prominent theoretical framework for explaining emotional responses to music posits *episodic memory* and *visual imagery* as separate mechanisms by which music induces emotions (Juslin, 2013). However, it is clear from previous research on MEAMs and episodic memory in general that there is much overlap between these two mechanisms. Indeed, it may be that *episodic memory* is more logically classified as a subcategory of *visual imagery*, if it is found to be the case that MEAMs always comprise some component of visual imagery. Similarly, future studies of the *visual imagery* mechanism should explore how often music-evoked visual imagery comprises scenes that could be classed as episodic memories (rather than, for instance, abstract shapes or imaginary scenes; Taruffi & Küssner, this volume).

Another important topic in this field is memory decline, as a result of both healthy ageing and disease or brain damage. Some research has revealed a decline in both autobiographical memory retrieval and visual imagery in Alzheimer's disease (AD; El Haj et al., 2016). However, several studies have provided evidence of relatively spared retrieval of autobiographical memories in AD when cued by music as compared to memories evoked in silence or via visual cues (e.g., Baird et al., 2018; El Haj et al., 2012b). Of particular relevance to the present focus, El Haj et al. (2012a, 2012b) found that MEAMs of people with AD were more specific (scored via the TEMPau test; Piolino et al., 2006) and had greater emotional impact than autobiographical memories generated in silence, although these MEAMs were still less specific than those of healthy control participants. Subsequent research is needed which further probes the particular content and features of the mental imagery underlying MEAMs in people with AD and other memory disorders. In addition, Belfi et al. (2018) have identified the mPFC as a crucial brain structure for vivid reliving of MEAMs, and Janata (2009) and Kubit and Janata (2018) have investigated the brain networks underlying MEAMs of healthy young adults. However, further exploration of the neurological underpinnings of MEAMs in people with memory impairments is needed to fully understand the conditions under which particular features of these memories may be spared.

In conclusion, mental imagery is a crucial component of autobiographical memory, which allows the rememberer to relive the sights, sounds, and other feelings of past experiences. Autobiographical memories evoked by music appear to be particularly vivid in several of these regards, which may be related to the remarkable frequency and diversity of ways with which people engage with music, and the particular value placed on music in developing and maintaining one's personal and social identity.

Notes

1 See also the work of Kubit and Janata (2018), who found that music-evoked autobiographical memories elicited activity in the default mode network, a network of brain regions typically implicated in introspection and mind-wandering.

2 Clarification in future studies is essential, since most previous studies have not differentiated between voluntary and involuntary MEAMs, and the extent to which retrieval intentionality might impact on the memory properties.

References

Aydin, C. (2018). The differential contributions of visual imagery constructs on autobiographical thinking. *Memory*, *26*(2), 189–200. https://doi.org/10.1080/09658211.2017.1340483

Baird, A., Brancatisano, O., Gelding, R., & Thompson, W. F. (2018). Characterization of music and photograph evoked autobiographical memories in people with Alzheimer's disease. *Journal of Alzheimer's Disease*, *66*(2), 693–706. https://doi.org/10.3233/JAD-180627

Belfi, A. M., Karlan, B., & Tranel, D. (2016). Music evokes vivid autobiographical memories. *Memory*, *24*(7), 979–989. https://doi.org/10.1080/09658211.2015.1061012

Belfi, A. M., Karlan, B., & Tranel, D. (2018). Damage to the medial prefrontal cortex impairs music-evoked autobiographical memories. *Psychomusicology: Music, Mind, and Brain*, *28*(4), 201–208. https://doi.org/10.1037/pmu0000222

Berntsen, D. (2009). *Involuntary autobiographical memories: An introduction to the unbidden past*. Cambridge: Cambridge University Press.

Berntsen, D., & Hall, N. M. (2004). The episodic nature of involuntary autobiographical memories. *Memory & Cognition*, *32*(5), 789–803. https://doi.org/10.3758/BF03195869

Brewer, W. F. (1986). What is autobiographical memory? In D. C. Rubin (Ed.), *Autobiographical memory* (pp. 25–49). New York: Cambridge University Press.

Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review*, *107*(2), 261–288. https://doi.org/10.1037/0033-295X.107.2.261

Crovitz, H. F., & Schiffman, H. (1974). Frequency of episodic memories as a function of their age. *Bulletin of the Psychonomic Society*, *4*(5), 517–518. https://doi.org/10.3758/BF03334277

Cuddy, L. L., Sikka, R., Silveira, K., Bai, S., & Vanstone, A. (2017). Music-evoked autobiographical memories (MEAMs) in Alzheimer disease: Evidence for a positivity effect. *Cogent Psychology*, *4*(1), 1277578. https://doi.org/10.1080/23311908.2016.1277578

El Haj, M., Fasotti, L., & Allain, P. (2012a). The involuntary nature of music-evoked autobiographical memories in Alzheimer's disease. *Consciousness and Cognition*, *21*(1), 238–246. https://doi.org/10.1016/j.concog.2011.12.005

El Haj, M., Kapogiannis, D., & Antoine, P. (2016). Phenomenological reliving and visual imagery during autobiographical recall in Alzheimer's disease. *Journal of Alzheimer's Disease*, *52*(2), 421–431. https://doi.org/10.3233/JAD-151122

El Haj, M., Postal, V., & Allain, P. (2012b). Music enhances autobiographical memory in mild Alzheimer's disease. *Educational Gerontology*, *38*(1), 30–41. https://doi.org/10.1080/03601277.2010.515897

Greenberg, D. L., & Rubin, D. C. (2003). The neuropsychology of autobiographical memory. *Cortex*, *39*(4–5), 687–728. https://doi.org/10.1016/S0010-9452(08)70860-8

Hassabis, D., & Maguire, E. A. (2007). Deconstructing episodic memory with construction. *Trends in Cognitive Sciences*, *11*(7), 299–306. https://doi.org/10.1016/j.tics.2007.05.001

Hyman, I. E., & Loftus, E. F. (1998). Errors in autobiographical memory. *Clinical Psychology Review*, *18*(8), 933–947. https://doi.org/10.1016/S0272-7358(98)00041-5

Jakubowski, K., Belfi, A. M., & Eerola, T. (2021). Phenomenological differences in music- and television-evoked autobiographical memories. *Music Perception: An Interdisciplinary Journal*, *38*(5), 435–455. https://doi.org/10.1525/mp.2021.38.5.435

Jakubowski, K., & Ghosh, A. (2019). Music-evoked autobiographical memories in everyday life. *Psychology of Music*. https://doi.org/10.1177/0305735619888803

Janata, P. (2009). The neural architecture of music-evoked autobiographical memories. *Cerebral Cortex*, *19*(11), 2579–2594. https://doi.org/10.1093/cercor/bhp008

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*(1), 54–75. https://doi.org/10.1037/a0024208

Janata, P., Tomic, S. T., & Rakowski, S. K. (2007). Characterisation of music-evoked autobiographical memories. *Memory*, *15*(8), 845–860. https://doi.org/10.1080/09658210701734593

Juslin, P. N. (2013). From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews*, *10*(3), 235–266. https://doi.org/10.1016/j.plrev.2013.05.008

Kristen-Antonow, S. (2019). The role of ToM in creating a reminiscence bump for MEAMs from adolescence. *Psychology of Music*, *41*(1), 51–68. https://doi.org/10.1177%2F0305735617735374

Kubit, B., & Janata, P. (2018). Listening for memories : Attentional focus dissociates functional brain networks engaged by memory-evoking music. *Psychomusicology: Music, Mind, and Brain*, *28*(2), 82–100. https://doi.org/10.1037/pmu0000210

Levine, B., Svoboda, E., Hay, J. F., Winocur, G., & Moscovitch, M. (2002). Aging and autobiographical memory: Dissociating episodic from semantic retrieval. *Psychology and Aging*, *17*(4), 677–689. https://doi.org/10.1037/0882-7974.17.4.677

Maes, P. J., Leman, M., Palmer, C., & Wanderley, M. M. (2014). Action-based effects on music perception. *Frontiers in Psychology*, *4*, 1–14. https://doi.org/10.3389/fpsyg.2013.01008

Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2003). *LIWC2001*. Mahwah: Lawrence Erlbaum Associates Inc.

Piolino, P., Desgranges, B., Clarys, D., Guillery-Girard, B., Taconnat, L., Isingrini, M., & Eustache, F. (2006). Autobiographical memory, autonoetic consciousness, and self-perspective in aging. *Psychology and Aging*, *21*(3), 510–525. https://doi.org/10.1037/0882-7974.21.3.510

Rentfrow, P. J., & Gosling, S. D. (2006). Message in a ballad: The role of music preferences in interpersonal perception. *Psychological Science*, *17*(3), 236–242. https://doi.org/10.1111/j.1467-9280.2006.01691.x

Rubin, D. C. (2005). A basic-systems approach to autobiographical memory. *Current Directions in Psychological Science*, *14*(2), 79–83. https://doi.org/10.1111%2Fj.0963-7214.2005.00339.x

Rubin, D. C., Burt, C. D. B., & Fifield, S. J. (2003b). Experimental manipulations of the phenomenology of memory. *Memory and Cognition*, *31*(6), 877–886. https://doi.org/10.3758/BF03196442

Rubin, D. C., Schrauf, R. W., & Greenberg, D. L. (2003a). Belief and recollection of autobiographical memories. *Memory and Cognition*, *31*(6), 887–901. https://doi.org/10.3758/BF03196443

Schulkind, M. D., & Woldorf, G. M. (2005). Emotional organization of autobiographical memory. *Memory and Cognition*, *33*(6), 1025–1035. https://doi.org/10.3758/BF03193210

Sheldon, S., Amaral, R., & Levine, B. (2017). Individual differences in visual imagery determine how event information is remembered. *Memory*, *25*(3), 360–369. https://doi.org/10.1080/09658211.2016.1178777

Sheldon, S., & Donahue, J. (2017). More than a feeling: Emotional cues impact the access and experience of autobiographical memories. *Memory and Cognition*, *45*(5), 731–744. https://doi.org/10.3758/s13421-017-0691-6

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology*, *26*(1), 1–12. https://doi.org/10.1037/h0080017

Uzer, T., & Brown, N. R. (2017). The effect of cue content on retrieval from autobiographical memory. *Acta Psychologica*, *172*, 84–91. https://doi.org/10.1016/j.actpsy.2016.11.012

Zator, K., & Katz, A. N. (2017). The language used in describing autobiographical memories prompted by life period visually presented verbal cues, event-specific visually presented verbal cues and short musical clips of popular music. *Memory*, *25*(6), 831–844. https://doi.org/10.1080/09658211.2016.1224353

12 What Is Mind-Wandering?

Mahiko Konishi

Imagine that you are driving home after a long day of work. You pass by the same buildings you see every day: the same intersection, the same supermarket, the same pharmacy. Suddenly, you are parking your car in your driveway; you do not remember much of the rest of the drive that got you there. What you do remember are the thoughts you had while you were on some sort of autopilot mode: the dinner you want to cook for yourself tonight; next week's important meeting; or fond memories of a good holiday, after an old summer hit comes up on the radio. It is a universal experience to drift off in one's own thoughts and worries; at times, these episodes can be so immersive that they make us lose touch with the external environment. Indeed, in many languages, the common name to refer to this experience is associated with dreaming: *daydreams* (English), *rêveries* (French), *soñar despierto* (Spanish), *sogni ad occhi aperti* (Italian), and *Tagtraum* (German).

This phenomenon has been long studied in experimental psychology under a range of different terminologies, such as "daydreaming," "task-unrelated thought," and "mind-wandering." These are just some of the words used to refer to *roughly* the same psychological experience. *Roughly* is an important adverb here: one of the most important, ongoing debates in the field of mind-wandering (MW) research revolves around defining what exactly should be classified as MW, and what shouldn't.

In fact, to date, there is no consensus definition of MW among the scientific community. Some researchers have tried to distinguish MW from other types of thought, using different criteria: for example, the thoughts experienced during MW should be unrelated to the primary task at hand (e.g., if I'm driving and thinking about driving, those thoughts would not count as MW; Smallwood & Schooler, 2015). Other researchers argue that MW should be free flowing, unconstrained, so that perseverative, ruminating thought such as common in major depressive disorder would not be categorized as MW (Christoff et al., 2018). Another perspective sees MW as a family of experiences that share some of their features with each other, but not necessarily having one in common among them all (Seli et al., 2018), borrowing the idea of *family resemblance* from Wittgenstein (1953/1972). This view also opens to the possibility that MW can be intentional or unintentional (Seli et al., 2016), while other camps would only define the latter as "true" MW. Agreement on the definition of MW, while difficult, remains one

DOI: 10.4324/9780429330070-16

of the most important open issues in the field, as a lack thereof hinders the comparison and evaluation of scientific results that used different definitions of this phenomenon.

While an exact definition still eludes researchers, what we do know is that MW is a ubiquitous experience. A few studies (Kane et al., 2007; Klinger & Cox, 1987; McVay et al., 2009) tried to estimate exactly how common this is, using an experience sampling approach (Csíkszentmihályi & Larson, 1987): this consists in briefly interrupting participants during their daily lives (e.g., with a beeper or a phone app) and asking them about the occurrence and the content of the thoughts they were having just before being interrupted. These studies showed that inter-individual rates of MW occurrence range enormously, with some participants reporting MW less than 10% of the time, while others more than 90% of the time; crucially, these experiments estimated that, on average, people spend between 25% and 30% of their waking lives engaged in MW.

In this chapter, I will try to answer some important questions regarding MW: Why do we spend so much time in our thoughts, and when are we most likely to do so? What do most people think about, when daydreaming? I will also look into some of the benefits that have been linked to MW, as well as its costs. Importantly, I will look into the relationship between MW and music, examining how music listening influences MW episodes (Herbert, this volume) and how MW can occur in the form of musical imagery (Floridou, this volume).

The Context and Contents of Mind-Wandering

Sitting on the train, looking out the window. Driving back home. Listening to a long meeting, late in the afternoon on a Friday. These are all situations in which we are likely to catch ourselves lost in our thoughts. Moreover, they illustrate the conditions that increase the chance of MW arising: a context in which the demands of the environment are low, or a state of boredom and fatigue. This intuition is backed up by psychological research. It has long been known that MW rates are higher when people are involved in easy tasks compared to harder ones (Antrobus et al., 1966; Konishi et al., 2017; Teasdale et al., 1995). Similarly, MW rates increase as a task is practiced more and more, effectively becoming easier to perform (Mason et al., 2007). Taking our driving example, it is unlikely that a beginner driver will MW much, focused as they are in the multitude of tasks that driving requires; on the other hand, an expert driver that has automatized this complex coordination of tasks will tend to daydream more, especially if driving along a well-known path such as the way back home from work. It is also known that MW increases together with fatigue and time-on-task, both in laboratory experiments (Smallwood et al., 2002) and in more ecological ones (Körber et al., 2015). As fatigue rises, we start to lose the ability to maintain sustained attention on one specific source (such as a monotonic speaker in a meeting), becoming more and more vulnerable to intrusions from our self-generated thoughts.

Several theories have been proposed to explain the occurrence of MW episodes: I here briefly summarize three prominent ones, taking the definitions from

Smallwood (2013). The *Current Concerns* hypothesis (Klinger & Cox, 2004) argues that MW occurs because the individual has goals and desires that extend beyond the present moment, and these can be chased through our self-generated thoughts, when the external environment offers less opportunities to chase such goals. The *Executive Failure* hypothesis (McVay & Kane, 2010) proposes that attention on the external environment is maintained through an executive system (an important part of the cognitive architecture, often invoked in cognitive psychology), which reduces distractions coming from both external sources and internal thoughts: in this view, MW occurrence reflects a failure of this system, allowing MW content to take over an individual's mental life (see also Gritten, this volume). Finally, the *Decoupling* hypothesis (Smallwood, 2013) sees internal and external processes as parallel trains of mental content; in this case, the executive system aims to maintain each train on its tracks when necessary, for example, by reducing daydreams when we have to focus on an external task (e.g., driving), or by reducing external distractions when we want to focus on our thoughts.

But what are our daydreams made of? Some studies have looked specifically into the contents of MW episodes. A typical finding in MW research, replicated with a range of different methodologies and in different cultures, is that people tend to think more about the future than the past (Baird et al., 2011; Iijima & Tanno, 2012). In other words, whenever our surroundings allow us to take a breather, we tend to think more about the future than the past (see Jakubowski, this volume, for an overview on research on autobiographical memories, imagery and music). Indeed, it is very efficient to be able to rehearse in our heads the presentation we will give in the afternoon, while sitting on the subway on our way to work.

As mentioned earlier, it is commonly found that a great deal of our daydreams seem centred around our personal goals. In fact, when dissected, most human experience and behaviour seem to be "organized around the pursuit and enjoyment of goals" (Klinger & Cox, 2004, p. 3). The idea that we constantly commit our behaviour, and our thoughts, around our *current concerns*, has been developed into a theory and has received empirical support (Klinger, 2009); following this theory, it has been hypothesized that most of MW reflects the activation of an individual's unresolved goals, expressed as daydreams (McVay & Kane, 2010; Smallwood & Schooler, 2006). For example, one might be daydreaming about the dinner they'll cook that night, because they haven't eaten since lunch, and as hunger creeps in, the goal of eating rises in priority among all their other goals.

We thus know a bit about what MW is made of, and the contexts in which it arises. But what are its consequences? Does MW have costs? Does it have benefits? The answer is yes, and yes.

The Benefits, and Costs, of Mind-Wandering

We have seen that MW is pervasive, taking a big chunk of our daily lives. This suggests that there might be benefits to this otherwise dangerous state of being decoupled from the external environment. As Antrobus et al. (1966, p. 399) put it: "The presence of this non-perceptual cognitive activity on such a large scale is perhaps the strongest argument that daydreaming and imagining serve a useful purpose for the individual." I already hinted to one crucial benefit before: as much of MW revolves around our personal goals (Medea et al., 2016; Stawarczyk et al., 2011), it allows us to prepare and plan for the future (Baird et al., 2011; Smallwood & Schooler, 2015). Indeed, the idea that mental time travel, such as conscious simulations of future events, is a major benefit for the organism has also been theorized in the field of self-regulation and emotional coping (Taylor & Schneider, 1989), and even as one of the evolutionary bases for the human mind (Suddendorf & Corballis, 1997). It has also been proposed that MW can facilitate and boost creativity, perhaps through a mechanism of unconscious incubation (Baird et al., 2012), helping to give rise to eureka (or "aha!") moments (Gable et al., 2019). It is worth noting that a few studies attempted to replicate some of these purported effects of MW on creativity, and failed to do so (Gardner, 2017; Smeekens & Kane, 2016), suggesting the fact that if these effects are real, they are likely to be weak or at least difficult to reproduce in a laboratory setting. There are indeed many analogies to be made between MW and creative thinking, and the two seem to share a neural substrate (Fox & Beaty, 2019).

The literature on the possible benefits of MW is in fact somehow limited. A few other benefits have been proposed (Mooneyham & Schooler, 2013), but so far, the available evidence points at the ability to prepare and plan for the future as the main benefit of MW. On the other hand, several costs of MW have been highlighted: MW has been linked to unhappiness and dysphoria (Smallwood & O'Connor, 2011), failures in reading comprehension (more on this will be discussed later), and decreases in driving performance (Baldwin et al., 2017; Yanko & Spalek, 2014), to attention and memory deficits (for an in-depth review, see Mooneyham & Schooler, 2013). We have all had the experience of reading a paragraph of a book again and again, only to realize we were thinking about something else entirely: indeed, experimental research confirmed that the frequency of MW during reading is closely related to decreases in comprehension, as well as changes in eye movements dynamics during reading (Reichle et al., 2010; Sanders et al., 2017; Unsworth & McMillan, 2012).

I have outlined how MW might have at least one crucial benefit, together with several accompanying costs. As I conclude this chapter, I will look at the relationship between MW and music.

Music and Mind-Wandering

How do music and MW interact? First, we can look at the effect of listening to music in evoking streams of thoughts and memories (Jakubowski, this volume), that is, how the occurrence and content of MW vary in a musical context. While this has been long known to happen anecdotally, recent research has started to formalize which features in the music, types of music, and listening settings can trigger these synaesthesia-like processes (e.g., Deil et al., 2022; Herff et al., 2021; Taruffi, 2021). For example, Taruffi et al. (2017) found that MW was more strongly associated with sad-sounding music relative to happy music, and that most of this type of music-evoked MW (also known as musical daydreaming, see Martarelli et al., 2016), happened in the form of visual mental images rather than in the linguistic form (like, e.g., an inner dialogue). Another study, this time comparing heroic and sad music, found evidence that listening to the former provoked more empowering and motivating thoughts, while listening to sad music produced more relaxing and sad thoughts (Koelsch et al., 2019). The phenomenology of this type of musical daydreaming is thoroughly investigated by Herbert (this volume).

Another way to look at the relationship between MW and music is to look at instances of musical imagery *within* MW episodes. Musical imagery generally refers to the experience of imagining music in our "mind's ear" (Floridou, this volume). This can be both voluntary, as when we start playing back in our head a song that we really like, or involuntary, when a catchy tune loops again and again in our thoughts against our will; this latter experience has also been dubbed *earworms* or *involuntary musical imagery* (INMI; Liikkanen, 2008). Like regular MW, musical imagery seems to be quite common: in two studies, participants reported imagining music in their heads 17% of the time they were probed (Bailes, 2015), with these rates almost doubling when involving music students (32%; Bailes, 2007). Similarly, another study surveying a large pool of participants found that 90% of them reported experiencing INMI at least once per week, and around one-third reporting at least one episode per day (Liikkanen, 2012).

As others have pointed out before (Farrugia et al., 2015; Liikkanen, 2008), INMI is a spontaneous cognitive phenomenon, and as such is closely related to MW. However, research on MW and on INMI has largely developed separately, on parallel tracks. While a few studies explored the two phenomena in unison (Floridou et al., 2017; Floridou & Müllensiefen, 2015), the great majority of MW studies using experience sampling did not include a question regarding musical imagery in their surveys. This has surely contributed to underestimating how common musical imagery is in our daily lives. Indeed, one study in our lab (Van den Driessche, et al., n.d.) found that, when participants were given the possibility of describing their mental content in an open-ended fashion, a non-trivial percentage of thoughts (~4%) could be classified as INMI; on the other hand, these would go unnoticed when participants had to rate their thoughts using standard categories in the MW field (e.g., on-task, MW, and distracted), commonly not including musical imagery categories.

The two fields have hence developed on parallel tracks, but an integrative approach seems fruitful: the field of INMI has investigated the phenomenology of this experience on features such as the duration of INMI episodes (for a general review, see Liikkanen, 2018), on individual differences (Beaty et al., 2013; Müllensiefen et al., 2014), and the potential triggers and causes of these episodes (Floridou & Müllensiefen, 2015). All these issues parallel unanswered questions in the MW field (Smallwood & Schooler, 2015). It thus seems undoubtedly advantageous for the field of MW to approach a phenomenon to whom it is closely related in terms of phenomenology and methodology, and whose research field has made advances that would benefit the MW field itself.

Mind-wandering is a familiar and common experience to us, and as such it has drawn the interest of experimental psychology and neuroscience. Both the complexity and the variety, intrinsic to the phenomenon, make it a difficult topic of scientific research. Nonetheless, several advances in understanding its cognitive and neural mechanisms have been recently made. As I have suggested, integrating its study with that of musical imagery seems a promising avenue to answer some of the outstanding questions in the field.

References

Antrobus, J. S., Singer, J. L., & Greenberg, S. (1966). Studies in the stream of consciousness: Experimental enhancement and suppression of spontaneous cognitive processes. *Perceptual and Motor Skills*, *23*(2), 399–417. https://doi.org/10.2466/pms.1966.23.2.399

Bailes, F. (2007). The prevalence and nature of imagined music in the everyday lives of music students. *Psychology of Music*, *35*(4), 555–570. https://doi.org/10.1177/0305735607077834

Bailes, F. (2015). Music in mind? An experience sampling study of what and when, towards an understanding of why. *Psychomusicology: Music, Mind, and Brain*, *25*(1), 58. http://dx.doi.org/10.1037/pmu0000078

Baird, B., Smallwood, J., Mrazek, M. D., Kam, J. W., Franklin, M. S., & Schooler, J. W. (2012). Inspired by distraction: Mind wandering facilitates creative incubation. *Psychological Science*, *23*(10), 1117–1122. https://doi.org/10.1177/0956797612446024

Baird, B., Smallwood, J., & Schooler, J. W. (2011). Back to the future: Autobiographical planning and the functionality of mind-wandering. *Consciousness and Cognition*, *20*(4), 1604–1611. https://doi.org/10.1016/j.concog.2011.08.007

Baldwin, C. L., Roberts, D. M., Barragan, D., Lee, J. D., Lerner, N., & Higgins, J. S. (2017). Detecting and quantifying mind wandering during simulated driving. *Frontiers in Human Neuroscience*, *11*, 406. https://doi.org/10.3389/fnhum.2017.00406

Beaty, R. E., Burgin, C. J., Nusbaum, E. C., Kwapil, T. R., Hodges, D. A., & Silvia, P. J. (2013). Music to the inner ears: Exploring individual differences in musical imagery. *Consciousness and Cognition*, *22*(4), 1163–1173. https://doi.org/10.1016/j.concog.2013.07.006

Christoff, K., Mills, C., Andrews-Hanna, J. R., Irving, Z. C., Thompson, E., Fox, K. C. R., & Kam, J. W. Y. (2018). Mind-wandering as a scientific concept: Cutting through the definitional haze. *Trends in Cognitive Sciences*, *22*(11), 957–959. https://doi.org/10.1016/j.tics.2018.07.004

Csíkszentmihályi, M., & Larson, R. (1987). Validity and reliability of the experience sampling method. *Journal of Nervous and Mental Disease*, *175*(9), 526–536. https://doi.org/10.1097/00005053-198709000-00004

Deil, J., Markert, N., Normand, P., Kammen, P., Küssner, M. B., & Taruffi, L. (2022). Mind-wandering during contemporary live music: An exploratory study. *Musicae Scientiae*. https://doi.org/10.1177/10298649221103210

Farrugia, N., Jakubowski, K., Cusack, R., & Stewart, L. (2015). Tunes stuck in your brain: The frequency and affective evaluation of involuntary musical imagery correlate with cortical structure. *Consciousness and Cognition*, *35*, 66–77. https://doi.org/10.1016/j.concog.2015.04.020

Floridou, G. A., & Müllensiefen, D. (2015). Environmental and mental conditions predicting the experience of involuntary musical imagery: An experience sampling method study. *Consciousness and Cognition*, *33*, 472–486. https://doi.org/10.1016/j.concog.2015.02.012

Floridou, G. A., Williamson, V. J., & Stewart, L. (2017). A novel indirect method for capturing involuntary musical imagery under varying cognitive load. *The Quarterly Journal of Experimental Psychology*, *70*(11), 2189–2199. https://doi.org/10.1080%2F17470218.2016.1227860

Fox, K. C., & Beaty, R. E. (2019). Mind-wandering as creative thinking: Neural, psychological, and theoretical considerations. *Current Opinion in Behavioral Sciences*, *27*, 123–130. https://doi.org/10.1016/j.cobeha.2018.10.009

Gable, S. L., Hopper, E. A., & Schooler, J. W. (2019). When the muses strike: Creative ideas of physicists and writers routinely occur during mind wandering. *Psychological Science*, *30*(3), 396–404. https://doi.org/10.1177/0956797618820626

Gardner, E. (2017). *Integrative mechanics of mind wandering and their effect on creativity* [Doctoral Dissertation, Haverford College]. Dspace Repositorium. https://scholarship.tricolib.brynmawr.edu/handle/10066/19415

Herff, S. A., Cecchetti, G., Taruffi, L., & Déguernel, K. (2021). Music influences vividness and content of imagined journeys in a directed visual imagery task. *Scientific Reports*, *11*(1), 1–12. https://doi.org/10.1038/s41598-021-95260-8

Iijima, Y., & Tanno, Y. (2012). The effect of cognitive load on the temporal focus of mind wandering. *Shinrigaku Kenkyu: The Japanese Journal of Psychology*, *83*(3), 232–236. https://doi.org/10.4992/jjpsy.83.232

Kane, M. J., Brown, L. H., McVay, J. C., Silvia, P. J., Myin-Germeys, I., & Kwapil, T. R. (2007). For whom the mind wanders, and when: An experience-sampling study of working memory and executive control in daily life. *Psychological Science : A Journal of the American Psychological Society/APS*, *18*(7), 614–621. https://doi.org/10.1111/j.1467-9280.2007.01948.x

Klinger, E. (2009). Daydreaming and fantasizing: Thought flow and motivation. In K. D. Markman, W. M. P. Klein, & J. A. Suhr (Eds.), *Handbook of imagination and mental simulation* (pp. 225–239). Hove: Psychology Press.

Klinger, E., & Cox, W. M. (1987). Dimensions of thought flow in everyday life. *Imagination, Cognition and Personality*, *7*(2), 105–128. https://doi.org/10.2190/7K24-G343-MTQW

Klinger, E., & Cox, W. M. (2004). Motivation and the theory of current concerns. In W. M. Cox & E. Klinger (Eds.), *Handbook of motivational counseling: Concepts, approaches, and assessment* (pp. 3–27). Hoboken: John Wiley & Sons Ltd.

Koelsch, S., Bashevkin, T., Kristensen, J., Tvedt, J., & Jentschke, S. (2019). Heroic music stimulates empowering thoughts during mind-wandering. *Scientific Reports*, *9*(1), 1–10. https://doi.org/10.1038/s41598-019-46266-w

Konishi, M., Brown, K., Battaglini, L., & Smallwood, J. (2017). When attention wanders: Pupillometric signatures of fluctuations in external attention. *Cognition*, *168*. https://doi.org/10.1016/j.cognition.2017.06.006

Körber, M., Cingel, A., Zimmermann, M., & Bengler, K. (2015). Vigilance decrement and passive fatigue caused by monotony in automated driving. *Procedia Manufacturing*, *3*, 2403–2409. https://doi.org/10.1016/j.promfg.2015.07.499

Liikkanen, L. A. (2008). Music in everymind: Commonality of involuntary musical imagery. In K. Miyazaki, Y. Hiraga, M. Adachi, Y. Nakajima, & M. Tsuzaki (Eds.), *Proceedings of the 10th International Conference on Music Perception and Cognition (ICMPC 10)* (pp. 408–412). Sapporo: ICMPC10.

Liikkanen, L. A. (2012). Musical activities predispose to involuntary musical imagery. *Psychology of Music*, *40*(2), 236–256. https://doi.org/10.1177/0305735611406578

Liikkanen, L. A. (2018). *Involuntary musical imagery: Everyday but ephemeral* [Doctoral Dissertation, Helsingin yliopisto]. HELDA. http://urn.fi/URN:ISBN:978-951-51-4408-9

Martarelli, C. S., Mayer, B., & Mast, F. W. (2016). Daydreams and trait affect: The role of the listener's state of mind in the emotional response to music. *Consciousness and Cognition*, *46*, 27–35. https://doi.org/10.1016/j.concog.2016.09.014

Mason, M. F., Norton, M. I., Van Horn, J. D., Wegner, D. M., Grafton, S. T., & Macrae, C. N. (2007). Wandering minds: The default network and stimulus-independent thought. *Science*, *315*(5810), 393–395. https://doi.org/10.1126/science.1131295

McVay, J. C., & Kane, M. J. (2010). Does mind wandering reflect executive function or executive failure? Comment on Smallwood and Schooler (2006) and Watkins (2008). *Psychological Bulletin*, *136*(2), 188–197. http://dx.doi.org/10.1037/a0018298

McVay, J. C., Kane, M. J., & Kwapil, T. R. (2009). Tracking the train of thought from the laboratory into everyday life: An experience-sampling study of mind wandering across controlled and ecological contexts. *Psychonomic Bulletin & Review*, *16*(5), 857–863. https://doi.org/10.3758/PBR.16.5.857

Medea, B., Karapanagiotidis, T., Konishi, M., Ottaviani, C., Margulies, D., Bernasconi, A., . . . Smallwood, J. (2016). How do we decide what to do? Resting-state connectivity patterns and components of self-generated thought linked to the development of more concrete personal goals. *Experimental Brain Research*, *236*(9), 2469–2481. https://doi.org/10.1007/s00221-016-4729-y

Mooneyham, B. W., & Schooler, J. W. (2013). The costs and benefits of mind-wandering: A review. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *67*(1), 11–18. https://doi.org/10.1037/a0031569

Müllensiefen, D., Fry, J., Jones, R., Jilka, S., Stewart, L., & Williamson, V. J. (2014). Individual differences predict patterns in spontaneous involuntary musical imagery. *Music Perception: An Interdisciplinary Journal*, *31*(4), 323–338. https://doi.org/10.1525/mp.2014.31.4.323

Reichle, E. D., Reineberg, A. E., & Schooler, J. W. (2010). Eye movements during mindless reading. *Psychological Science*, *21*(9), 1300–1310. https://doi.org/10.1177/0956797610378686

Sanders, J., Wang, H.-T., Schooler, J., & Smallwood, J. (2017). Can I get me out of my head? Exploring strategies for controlling the self-referential aspects of the mind-wandering state during reading. *Quarterly Journal of Experimental Psychology*, *70*(6), 1053–1062. https://doi.org/10.1080/17470218.2016.1216573

Seli, P., Kane, M. J., Smallwood, J., Schacter, D. L., Maillet, D., Schooler, J. W., & Smilek, D. (2018). Mind-wandering as a natural kind: A family-resemblances view. *Trends in Cognitive Sciences*, *22*(6), 479–490. https://doi.org/10.1016/j.tics.2018.03.010

Seli, P., Risko, E. F., & Smilek, D. (2016). On the necessity of distinguishing between unintentional and intentional mind wandering. *Psychological Science*, *27*(5), 685–691. https://doi.org/10.1177/0956797616634068

Smallwood, J. (2013). Distinguishing how from why the mind wanders: A process: Occurrence framework for self-generated mental activity. *Psychological Bulletin*, *139*(3), 519. https://psycnet.apa.org/doi/10.1037/a0030010

Smallwood, J., Obonsawin, M., & Reid, H. (2002). The effects of block duration and task demands on the experience of task unrelated thought. *Imagination, Cognition and Personality*, *22*(1), 13–31. https://doi.org/10.2190/TBML-N8JN-W5YB-4L9R

Smallwood, J., & O'Connor, R. C. (2011). Imprisoned by the past: Unhappy moods lead to a retrospective bias to mind wandering. *Cognition & Emotion*, *25*(8), 1481–1490. https://doi.org/10.1080/02699931.2010.545263

Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin*, *132*(6), 946–958. https://doi.org/10.1037/0033-2909.132.6.946

Smallwood, J., & Schooler, J. W. (2015). The science of mind wandering: Empirically navigating the stream of consciousness. *Annual Review of Psychology*, *66*(1), 487–518. https://doi.org/10.1146/annurev-psych-010814-015331

Smeekens, B. A., & Kane, M. J. (2016). Working memory capacity, mind wandering, and creative cognition: An individual-differences investigation into the benefits of controlled versus spontaneous thought. *Psychology of Aesthetics, Creativity, and the Arts*, *10*(4), 389. http://dx.doi.org/10.1037/aca0000046

Stawarczyk, D., Majerus, S., Maj, M., Van der Linden, M., & D'Argembeau, A. (2011). Mind-wandering: Phenomenology and function as assessed with a novel experience sampling method. *Acta Psychologica*, *136*(3), 370–381. https://doi.org/10.1016/j.actpsy.2011.01.002

Suddendorf, T., & Corballis, M. C. (1997). Mental time travel and the evolution of the human mind. *Genetic, Social, and General Psychology Monographs*, *123*(2), 133–167.

Taruffi, L. (2021). Mind-wandering during personal music listening in everyday life: Music-evoked emotions predict thought valence. *International Journal of Environmental Research and Public Health*, *18*(23), 12321. https://doi.org/10.3390/ijerph182312321

Taruffi, L., Pehrs, C., Skouras, S., & Koelsch, S. (2017). Effects of sad and happy music on mind-wandering and the default mode network. *Scientific Reports*, *7*(1), 14396. https://doi.org/10.1038/s41598-017-14849-0

Taylor, S. E., & Schneider, S. K. (1989). Coping and the simulation of events. *Social Cognition*, *7*(2), 174–194. https://doi.org/10.1521/soco.1989.7.2.174

Teasdale, J. D., Dritschel, B. H., Taylor, M. J., Proctor, L., Lloyd, C. A., Nimmo-Smith, I., & Baddeley, A. D. (1995). Stimulus-independent thought depends on central executive resources. *Memory & Cognition*, *23*(5), 551–559. https://doi.org/10.3758/BF03197257

Unsworth, N., & McMillan, B. D. (2012). Mind wandering and reading comprehension: Examining the roles of working memory capacity, interest, motivation, and topic experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(3), 832–842. https://doi.org/10.1037/a0029669

Van den Driessche, C., Chappé, C., Konishi, M., & Sackur, J. (n.d.). *Rapports verbaux des contenus mentaux et validation de la classification* [Unpublished Manuscript]. Départment d'Études Cognitives, École Normale Supérieure.

Wittgenstein, L. (1972). *Philosophical investigations* (G. E. M. Anscombe, Trans., Rev. 3rd ed.). Oxford, UK: Basil Blackwell. (Original work published in 1953)

Yanko, M. R., & Spalek, T. M. (2014). Driving with the wandering mind: The effect that mind-wandering has on driving performance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *56*(2), 260–269. https://doi.org/10.1177/0018720813495280

13 Distraction or Panic?

Anthony Gritten

This chapter concerns a phenomenon that disrupts mind-wandering while accounting for its spontaneity: distraction. It is about one of the mechanisms underlying mind-wandering rather than its phenomenological content. However, given that mind-wandering has positive and negative consequences for task fulfilment (Konishi, this volume; Mooneyham & Schooler, 2013), this chapter recuperates distraction, configuring it as "drag" in order to analyse it as more than merely bad or failed listening. Dragging the listener back onto her body, distraction emphasizes indeterminacies that, although covered over by the traction generated by regular and consistent listening, are always already embodied within the materiality of musical sound. This chapter concerns the extent to which distraction forces the listener to stretch herself between "listening despite distraction" and "distracted listening"; a separate chapter would be required to consider the different ways that performers confront distraction.

Distraction

On August 2, 1994, the BBC Symphony Orchestra and Oliver Knussen gave the UK premiere of Alexander Goehr's *Colossos or Panic: Symphonic Fragment after Goya* op. 55 (1991–92) at the Proms. This performance, which I attended with my father, characterized the dilemma facing contemporary listeners: between configuring distraction as an event that disrupts listening and configuring it as a necessary and welcome component of attentive listening. Calling this a "dilemma," suggesting choice and decision, is misleading, but it emphasizes the cultural register of distraction and its imbrication with mind-wandering.

During what I subsequently came to know as the last few pages of the first movement, there was a disruptive event: a helicopter flying somewhere overhead. Everybody in the Albert Hall heard it, and the identity of the sound was unmistakeable. It lasted long enough for some listeners to begin murmuring in disquiet, while others chuckled. On first hearing, the helicopter noise seemed to distract from the performance, for three reasons: it was louder than the music; it had not been planned; and it could not be assimilated into the ongoing performance. The helicopter noise turned my mind away temporarily from what I had been doing quietly for the first 12 minutes of the performance: listening to the music, my

DOI: 10.4324/9780429330070-17

mindwandering off now and then as I pondered the relation between the unfolding musical flow and the Goya painting that had set Goehr in motion.

The event was more complex than this, however. As the helicopter flew away, a temporary feeling of surprise pervaded the hall as it became clear that the sound around Figure [108] in the subsequently published score had been both the helicopter *and* a low sustained E natural in the double basses and bassoons. The one had become the other, temporarily, and the quasi-tonal function of the noise *qua* music seemed both to extend outwards from the orchestra and, in reverse, to insert itself into the orchestra from outside, in an uncannily immersive moment of music *qua* noise. In the second movement, too, some low pedal points echoed the deep pitch of the now-absent helicopter, their quasi-dominant tonal pull creating a similar aesthetic expectation of musical movement that the external noise had generated at the end of the first movement. Given that the "pitch" and profile (the Doppler effect) of the helicopter seemed therefore, to my naïve ears, to fit with the post-tonal harmonic profile of the end of the first movement, providing resolution to the musical activity prior to this point, I now ask, what constitutes a classical musical sound within an urban environment? This question is not trivial. It is equivalent to asking what constitutes distraction during a performance of a work such as Goehr's, and how much energy is required to recognize and situate a classical musical sound within the ongoing development of auditory traction. Separating musical sound from noise is predicated upon processing noise during music, the two registers of auditory activity—sound and noise—being intimately related. I am not suggesting that the helicopter was not a distraction from what we had come to hear that evening. It was, and, had we been listening at home to the live radio broadcast, I might have found these questions harder to disentangle without the evolving acoustic context available in the Albert Hall. Distractions are contextual, emerging relative to auditory streams into which they insert themselves (the world premiere of *Colossos or Panic*, available on YouTube (Goehr, 1993b), contained no distraction at this point).

Rather, the challenge to contemporary Western classical music—"Musik unserer Zeit," says Goehr's publisher—presented by the urban environment is split between managing the tendency towards distracted listening and managing listening in a sustainable manner *despite* distraction. The first task is particularly pressing when, like *Colossos or Panic*, the music is complex modernist symphonic music in the Western classical tradition predicated upon attentive listening; substituting Radiohead for Goehr—"In the city of the future/It is difficult to concentrate" (Radiohead, 1998)—generates different questions. The second part of the challenge is made harder by the classical recording industry, which for a century has been assiduously removing distracting noises from recordings and thus preparing listeners increasingly badly for worldly listening. Listeners of Western classical music still often assume that the urban environment is distracting them from the aesthetic world. It would require separate essays to ask *how* Western classical music evolved this way, and *why* listening distractedly has not caught on with *Colossos or Panic* as it has with albums like *Airbag/How am I driving?* (Radiohead, 1998). Goehr (1998) himself has explored such questions.

While more broadly it has been argued that, historically, "Attention and distraction were not two essentially different states but existed on a single continuum" (Crary, 1999, p. 47), it is just as pressing to explore this claim regarding today's distracted world, so my concern with the helicopter at *Colossos or Panic*'s premiere is with the complex relationship between traction and distraction. How can we engage this event of distraction, given the porous boundary between musical sound and noise?

Panic?

We begin with the listening regimes that emerge on the back of traction. I take traction to be the ability to follow music reliably and recognize relevant connections between sounds as they emerge during live performance.

Listening regimes offer "integration strategies" (Bigand et al., 2000, p. 277) that regulate access to musical sound, filter perceptions, and focus the ears. They are essentially rhythmic practices, rules configuring the pace, intensity, and structural characteristics of the listener's engagement with sound, and affording it coherence, shape, and focus. Providing a *telos* for the listener's energetic expenditure and "listening effort" (Fairnie et al., 2016, p. 937), listening regimes include structural listening, pedagogical listening, hermeneutic listening, therapeutic listening, purposeless listening (Cage, 1961), concatenated listening (Levinson, 1997), self-aware listening (Schäfer et al., 2013), and political listening (Waltham-Smith, 2017). Regimes have two characteristics: psychologically, the listener is brought closer to the music; phenomenologically, feelings of purity, pleasure, and wholeness increase as the listener masters the regime.

However, it was clear from *Colossos or Panic* that listening has a complex relationship to the materiality of musical sound. Musical sound is embodied in energetic disturbances of air and material vibrations of the body, and, notwithstanding the tendency for listening regimes to emphasize aesthetic registers of traction, material vibrations are restricted neither to metaphorical registers nor merely to the ears: vibrations affect the entire body, which resonates with the disturbance that is musical sound, particularly in live performances where the acoustic properties of halls play a large role. Thus, while being a reliable listener and developing a degree of traction on musical sound means that the listener can feel some degree of ownership of the musical sound (listening again tomorrow will be broadly similar), "ownership" is an unhelpful term. For musical sound is not bound by the demands of the listener developing traction. Its trajectory is a function of energetic expenditure and of entropy, and its temporal envelope—during which the listener actively and sensuously engages it—is governed by the eventual dispersal of its vibrations into the surrounding environment. Sound is "an experience without transcendental guarantees" (North, 2011, p. 217, note 13).

The relationship, then, between musical sound and its listener is more complex than reciprocity, and the fact that it is often phrased the other way around (me and *my* listening) is unhelpful. It is less that, through traction, the listener owns the musical sound, more that she is loaned the possibility of traction. While classical phenomenology (Husserl, 1928/1991) argues that developing auditory traction is a matter of pursuing musical sound into the future and negotiating a deal between subject and object, what transpires is a compromise. Musical sound is forever loaning traction with one hand and taking it away with the other, in a proportion depending on the particular musical style and the acoustic environment. No amount of energetic expenditure guarantees full ownership and control of musical sound, even when such things are figured into the aesthetic of the music, as with structural listening and modernist music. The listener's relationship to musical sound moves down the slope of energetic expenditure into "fatigue" and "inattentional deafness" (Scheer et al., 2018, p. 429). Traction develops at the price of its dispersal by musical sound. Given that, "projecting a motor intentionality toward musical events, the body is *already* doing analytical work" (Kozak, 2015, p. 9), it is clear that the materiality of musical sound is intimately related to the listener's own materiality. Thus, open both to the wonder of musical sound's emergence and to its desperately bad timing (the listener's and musical sound's temporalities head in different directions), listening develops traction on musical sound while the sound disperses according to its own internal temporality (hence, memory is important to understanding). If musical sound seems to confound the senses, to set the ears in motion but itself not last long enough to hear what they generate by way of traction (which extends beyond the acoustic end of the note into retention and memory), then a fully developed listening would need, not just to develop traction each time it engages musical sound, but to build on each temporary achievement and acknowledge its own indeterminacy and unpredictability, the fact that "sensory perception cannot be separated from the multisensory experience of our bodies" (Phillips-Silver & Trainor, 2007, p. 543).

The parallel event streams of musical sound reify into percepts only slightly more permanent than those demanded by local auditory binding and the generation of traction. Musical sound presents listening with resonant materiality. Given that a large part of musical sound's materiality functions largely before listening arrives on the auditory scene, and at a non-conscious level (Dykstra et al., 2017, p. 13), how can the listener recognize it? One way involves focusing on moments when listening falters and adapts, when there is "a brief degradation of performance" (Rinne et al., 2007, p. 247), as with the swerve in my listening to Goehr. Being subject to the entropic materiality of musical sound, listening is stretched between indeterminacy and traction, and what classical phenomenology describes as the "adequation" of act and object needs revision as follows: musical sound disperses the listening subject around resonant space. It reminds the listener of her indeterminacy and susceptibility to sound, that musical sound is uncontainable. Listening draws something out of materiality against the odds, responding to musical sound itself rather than the proprietary fantasies of a listening regime. In this respect, the materiality of musical sound ensures the listener's awareness of her embodiment within the sound world.

With this description of musical sound's materiality in mind, we return to the complexity of my embodied perception of orchestral timbre in the Albert Hall. At issue is how I felt that specific event of "distraction within listening." I deploy this phrase rather than "distracted listening," because my focus is less on listening that attempts to accept distraction (whether a lifestyle choice, facilitated by digital media, or lack of time), and more on listening that attempts to minimize it in order to focus on the aesthetic register of attentive listening. The two modes of listening are not mutually exclusive, and Goehr's listener today needs to be pragmatic about balancing them together, but my focus here is on the fate of listening *despite* distraction, since it is this that characterized the aesthetic ideology of the particular sub-culture of contemporary Western classical music within which *Colossos or Panic* was situated. A broader essay could consider *why* a listener might attempt to engage or reject distraction in a particular way; here I consider the underlying dynamics of *how* my listening engaged with the sound dispersing around the Albert Hall, given that the music generated by the orchestral musicians was a tiny subset of sound in general, emerging within intentionality as much through exclusion as through inclusion.

The relationship between distraction and attentive listening is more complex than mere opposition (Benjamin, 1968). Distraction can impede listening, embody a release from listening, interrupt listening, intensify listening, and more. Some relationships are positive, others are negative; and they can be phrased with passive grammar—listening can be impeded by distraction—although this situates agency quite differently. Cutting across simplistic capitalist distinctions between distractions that help productivity and distractions that hinder it, there are two main types of distraction: "The first—*competition for action*—occurs when the results of obligatory musical sound processing are similar to those of the focal task. The second—*interruption of action*—takes place when an unexpected musical sound draws attention away from the focal activity" (Hughes et al., 2011, p. 1). Cognitive models of distraction configure it in three stages: perceiving change, orienting to the distraction, and re-orienting to the primary task.

Unpacking distraction in terms of such relationships, I displace concern for efficacious, efficient, and effective listening with phenomenological issues of "competition" and "interruption." Distraction drags the listener's body back into the world of energetic expenditure. It is cross-modal (Hubbard, 2018), and interferes with "several functional systems" (Rinne et al., 2007, p. 250). Thus, dragged listening is a matter of enactive cognition (Palmiero et al., 2019) that adapts to the disruptive event and recalibrates the listening body relative to the musical sound. It is unnecessary to assume that attentive listening without or prior to distraction is disembodied, pure, and efficient; attention itself "is not a unitary phenomenon" (Anderson, 2016; Anderson & Druker, 2013). For the purposes of understanding listening as proprioceptive (Acitores, 2011), however, it is useful heuristically to posit two tendencies, one towards greater embodiment, the other towards greater imaginative representation, in order to show that distraction is more complex than the issue of "divided attention" grounding music perception (Bigand et al., 2000, p. 277). The issue here is the variable ownership of listening, the fact that attentive listening is not situated contra distraction, but that it is always already contaminated by distraction.

We might say that "traction engines" are tools extending the listening body while "distraction machines" are autonomous mechanisms programmed to constitute the listening body. Traction is subject to time, while distraction is a "technics" of time. Traction develops order in the evolving relation with musical sound, while distraction complexifies this order by injecting indeterminacy into it. Traction is a trace of the body's attempt to accumulate time, to model timespans rhythmically, while distraction disperses the listening body's accumulated time, and forces the listener to start another timespan. Traction brings the body together under a single umbrella, that of the seductive "will-to-synchronize" (Pettman, 2016, p. 123) that dominates the distracted contemporary world, while distraction disperses it around resonant space, sometimes unbeknownst to the listener.

Given that intentionality is the ground of attention (e.g., Husserl, 1928/1991), it is less that distraction replaces traction, more that it displaces it with directives to re-embody listening. I probably overemphasize the disruptive energy of distraction's drag on the listener and its intervention into the embodiment of musical sound, the indeterminacy of musical sound in a state of distraction, but distraction's drag is also a potential source of traction in listening. Indeed, acknowledging that listening is constituted simultaneously by traction and distraction, not in circular but in spiralling mutually triggering events, raises the issue of when reliability is better than indeterminacy, and vice versa; the answer, of course, is that it varies according to local aesthetic context. Listening to *Colossos or Panic*, I had images of the orchestral music *qua* noise juxtaposed with images of the helicopter noise *qua* music. My mind wandered into a hybrid space in which the phenomenological *qualia* of the two events became indiscernible, and in which other aspects of the music—the diffuse chamber texture around Figure [108] in the published score, the falling trumpet F#-E#-C#—ceased to generate *musical* traction and the indeterminacy of the distraction covered over my perception until the source of the disturbance had flown out of earshot.

In such moments, distraction is an obstacle in the path of traction and attentive listening, and, notwithstanding my ability to undertake notational audiation, such moments drift from my grasp. Although auditory imagery in the absence of stimuli involves the same brain areas as auditory perception (Belfi, this volume; Hubbard, 2010), distraction from *Colossos or Panic* subsumed the empirical register of musical sound and overwhelmed its imaginative register, turning my feeling of traction back onto itself, dragging my listening off track and into the unfolding flow. It caused unintentional mind-wandering (Seli et al., 2016) where my off-task thoughts took time to reorientate themselves with respect to the primary task. And while the loss of traction was only temporary, the distraction event was more than a minor event that could have dispersed with better traction (if I had known the music), for dispersing musical time and emerging between one event and the next, distraction causes a temporary distortion in the flow of events, "a diaspora without hope or promise" (North, 2011, p. 146). Using its force to insinuate an unplanned change of direction, its intervention into the listener's developing traction directs attention towards what was not otherwise happening.

Regarding the four theories of mind-wandering—perceptual decoupling, executive failure, the current concerns hypothesis, and resource control accounts (Konishi, this volume)—some basic commonalities can be noted to distraction. What is distracting in one context may not be as distracting elsewhere (Chen & Sussman, 2013). Distractions do not have to be loud (Schröger et al., 2000). The degree of distraction is a function of "the particular processes engaged, rather than the amount or complexity of information being processed" (Hemond et al., 2010, p. 652). Different types of cognitive load impact differently upon the likelihood of distraction, for example, "perceptual load and working memory load have opposite effects on selective attention" (Lavie et al., 2004, p. 348). Suppressing information irrelevant to the task occurs before conflicts are resolved between perceptual information (Stewart et al., 2017). The complexity of interruptions influences their disruptive potential (Cades et al., 2008). The dichotic listening and irrelevant sound paradigms deal with "unattended sound" in different ways (Beaman et al., 2007). Distractions that are semantically related to the contents of primary perception may interfere with memory (Beaman et al., 2013; Marsh et al., 2013).

What distraction does to accumulated traction is fold the body onto itself, while the listening *of* the listener remains supervenient on her listening *to* musical sound. The listener's development of traction on sound and subscription to an appropriate listening regime is dragged into indeterminacy and delayed. The energetic expenditure of dragging the listener back onto her body embodies a further temporal stream, loosening the intimacy of the phenomenological *noesis-noema* and letting musical sound disperse in a temporary "dissolution" of her faculties (North, 2011, p. 15). As one study of auditory load concludes:

> [I]t seems that there is a finite auditory perceptual capacity that is assigned in an automatic fashion until resources are exhausted. Additional processing of distractors, or secondary task performance, therefore depends on whether any spare capacity remains after resources are assigned to task-relevant processing.
>
> (Fairnie et al., 2016, p. 935)

Distraction's lesson is that traction is open to adaptation and compromise—as the listener quickly *feels* when the continuity of a listening regime is disrupted. In my helicopter experience, distraction set in motion a brief moment of mind-wandering away from the sensuous immediacy of the unfolding musical sound, which followed large-scale alternations between chamber groupings centred on oboe, clarinet, and violin solos and increasingly intense waves of repeating *tutti* material, all climaxing between Figures [100] and [107]. As the music dissolved into quieter gestures, I found myself thinking about the larger sonic environment within which the concert was happening and about quite how permeable modernist symphonic sound is to what is usually denigrated as "noise." Regarding theories of mind-wandering (Christoff et al., 2016), these may have been task-unrelated thoughts; they were also spontaneous and constrained—directed by the sheer materiality of the helicopter's auditory presence. Thinking about what was happening while it happened did not help, since consciousness can "wander when we try to hold it in place while we check to make sure it has not moved" (Wegner, 1997, p. 312). Ultimately, the ongoing musical flow into the start of the second movement put my listening back on track.

In this respect, distraction's drag seems to provide listening with a rhetoric of embodiment. The discourse of living in the distracted contemporary world predicts this conclusion. However, distraction is cunning: "In becoming a habit, distraction becomes a tool for dissolving regimes of thought, modes of understanding, by admitting an empirical moment unto the transcendental structure of apperception" (North, 2011, pp. 164–165). It undoes not only traction's claim to have grasped musical sound's materiality but also its own claim to be a force for perceptual adaptation. A distraction fully cognized finds its *qualia* changed (no longer distracting) and become something for the listening body. Such is how, a quarter of a century later, I understand that experience in the Albert Hall. This does not mean that distraction should trigger panic, but simply that the listener's task is to engage musical sounds on the basis of drag: in line with their physical embodiment, vibration, resonance, and the mind-wandering that they both set in motion and disrupt. She must adapt to what happens and acknowledge calmly that musical sound has the last word—which is not a word but the silence of its entropic dispersal.

Conclusion

My argument recuperates positivity from distraction's infamous negativity. I have described face-to-face attentive listening—the backbone of my conventional education—as it functioned and then failed temporarily, my brain responding to the distraction event in the urban environment that disrupted the premiere of *Colossos or Panic*. My conclusion is that distraction may be a drag on the listener, and an event that complexifies mind-wandering, but its force for positive change—challenging the listener to adapt to the materiality of musical sound—should be acknowledged in our accounts of listening, whether to Radiohead or to Goehr. Primarily, the listener must develop the ability to adapt pragmatically to what happens in the environment, and distraction's retention of an essential indeterminacy within listening affords her precisely this ability. Ultimately, attempting to listen *despite* distraction is a mistaken expenditure of energy, and not entirely pragmatic, given the noisy soundscapes of contemporary society.

I have dragged *Colossos or Panic* into my argument, which may have been unfair. Goehr's concerns are not mine. Yet this work was composed after the fall of the Berlin wall when there were prospects of global peace and promises—threats?—of neoliberal capitalism were on the horizon. Just as Goya's 1808–12 painting can be read allegorically, so can Goehr's music. My reading, hiding behind noisy political questions, argues that *Colossos or Panic* embodied a pressing issue concerning the function of attentive listening in the distracted contemporary world: whether it was still possible to resist, minimize, or simply live alongside distraction. I close with the wisdom from Goehr's text on Goya (1993a, p. 25): "We may imagine that, like a rainbow, it will of a sudden be gone; all will be as before."

References

Acitores, A. P. (2011). Towards a theory of proprioception as a bodily basis for consciousness in music. In D. Clarke & E. Clarke (Eds.), *Music and consciousness: Philosophical, psychological, and cultural perspectives* (pp. 215–230). Oxford: Oxford University Press.

Anderson, B. (2016). Attentional effects on phenomenological appearance: How they change with task instructions and measurement methods. *PLoS One*, *11*(3), e0152353. https://doi.org/10.1371/journal.pone.0152353

Anderson, B., & Druker, M. (2013). Attention improves perceptual quality. *Psychonomic Bulletin and Review*, *20*(1), 120–127. https://doi.org/10.3758/s13423-012-0323-x

Beaman, C. P., Bridges, A. M., & Scott, S. K. (2007). From dichotic listening to the irrelevant sound effect: A behavioural and neuroimaging analysis of the processing of unattended speech. *Cortex*, *43*(1), 124–134. https://doi.org/10.1016/S0010-9452(08)70450-7

Beaman, C. P., Hanczakowski, M., Hodgetts, H. M., Marsh, J. E., & Jones, D. M. (2013). Memory as discrimination: What distraction reveals. *Memory & Cognition*, *41*(8), 1238–1251. https://doi.org/10.3758/s13421-013-0327-4

Benjamin, W. (1968). The work of art in the age of mechanical reproduction (H. Zohn, Trans.). In W. Benjamin (Ed.), *Illuminations: Essays and reflections* (pp. 217–252). New York: Schocken.

Bigand, E., McAdams, S., & Forêt, S. (2000). Divided attention in music. *International Journal of Psychology*, *35*(6), 270–278.

Cades, D. M., Werner, N., Boehm-Davis, D. A., Trafton, J. G., & Monk, C. A. (2008, September). Dealing with interruptions can be complex, but does interruption complexity matter: A mental resources approach to quantifying disruptions. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 52, No. 4, pp. 398–402). Los Angeles: Sage Publications. https://doi.org/10.1177%2F154193120805200442

Cage, J. (1961). *Silence: Lectures and writings*. Middletown: Wesleyan University Press.

Chen, S., & Sussman, E. S. (2013). Context effects on auditory distraction. *Biological Psychology*, *94*(2), 297–309. https://doi.org/10.1016/j.biopsycho.2013.07.005

Christoff, K., Irving, Z. C., Fox, K. C., Spreng, R. N., & Andrews-Hanna, J. R. (2016). Mind-wandering as spontaneous thought: A dynamic framework. *Nature Reviews Neuroscience*, *17*(11), 718–731. https://doi.org/10.1038/nrn.2016.113

Crary, J. (1999). *Suspensions of perception: Attention, spectacle, and modern culture*. Cambridge: MIT Press.

Dykstra, A. R., Cariani, P. A., & Gutschalk, A. (2017). A roadmap for the study of conscious audition and its neural basis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1714), 20160103. https://doi.org/10.1098/rstb.2016.0103

Fairnie, J., Moore, B. C., & Remington, A. (2016). Missing a trick: Auditory load modulates conscious awareness in audition. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(7), 930–938. http://dx.doi.org/10.1037/xhp0000204

Goehr, A. (1998). *Finding the key: Selected writings of Alexander Goehr*. London: Faber & Faber.

Goehr, A. (1993a). Programme note for *Colossos or Panic: Symphonic Fragment after Goya* op. 55. In *Boston Symphony Orchestra programme for 112th season 1992–93* (pp. 23–25). Boston: BSO.

Goehr, A. (1993b). *Colossos or Panic: Symphonic Fragment after Goya op. 55*. Premiere performance by Boston Symphony Orchestra and Seiji Ozawa. https://youtu.be/6QD0-iv55ww

Hemond, C., Brown, R. M., & Robertson, E. M. (2010). A distraction can impair or enhance motor performance. *Journal of Neuroscience*, *30*(2), 650–654. https://doi.org/10.1523/JNEUROSCI.4592-09.2010

Hubbard, T. L. (2010). Auditory imagery: Empirical findings. *Psychological Bulletin*, *136*(2), 302. https://doi.org/10.1037/a0018436

Hubbard, T. L. (2018). Some methodological and conceptual considerations in studies of auditory imagery. *Auditory Perception & Cognition*, *1*(1–2), 6–41. https://doi.org/10.1080/25742442.2018.1499001

Hughes, R. W., Vachon, F., Hurlstone, M., Marsh, J. E., Macken, W. J., & Jones, D. M. (2011, July). *Disruption of cognitive performance by sound: Differentiating two forms of auditory distraction* [Paper presentation]. 11th International Congress on Noise as a Public Health Problem (ICBEN), London, UK.

Husserl, E. (1991). *On the phenomenology of the consciousness of internal time (1893–1917)* (J. B. Brough, Trans.). Dordrecht: Kluwer Academic Publishers. (Original work published 1928)

Kozak, M. (2015). Listeners' bodies in music analysis: Gestures, motor intentionality, and models. *Music Theory Online*, *21*(3). www.mtosmt.org/issues/mto.15.21.3/mto.15.21.3.kozak.php

Lavie, N., Hirst, A., De Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, *133*(3), 339–354. https://doi.org/10.1037/0096-3445.133.3.339

Levinson, J. (1997). *Music in the moment*. Ithaca, New York: Cornell University Press.

Marsh, J., Sörqvist, P., Beaman, P., & Jones, D. (2013). Auditory distraction eliminates retrieval induced forgetting: Implications for the processing of unattended sound. *Experimental Psychology*, *60*(5), 368–375. https://doi.org/10.1027/1618-3169/a000210

Mooneyham, B. W., & Schooler, J. W. (2013). The costs and benefits of mind-wandering: A review. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *67*(1), 11–18. https://doi.org/10.1037/a0031569

North, P. (2011). *The problem of distraction*. Stanford: Stanford University Press.

Palmiero, M., Piccardi, L., Giancola, M., Nori, R., D'Amico, S., & Belardinelli, M. O. (2019). The format of mental imagery: From a critical review to an integrated embodied representation approach. *Cognitive Processing*, *20*(3), 277–289. https://doi.org/10.1007/s10339-019-00908-z

Pettman, D. (2016). *Infinite distraction: Paying attention to social media*. Cambridge: Polity.

Phillips-Silver, J., & Trainor, L. J. (2007). Hearing what the body feels: Auditory encoding of rhythmic movement. *Cognition*, *105*(3), 533–546. https://doi.org/10.1016/j.cognition.2006.11.006

Radiohead. (1998). Palo Alto [Song]. On *Airbag/How am I driving?*. Capitol.

Rinne, T., Kirjavainen, S., Salonen, O., Degerman, A., Kang, X., Woods, D. L., & Alho, K. (2007). Distributed cortical networks for focused auditory attention and distraction. *Neuroscience Letters*, *416*(3), 247–251. https://doi.org/10.1016/j.neulet.2007.01.077

Schäfer, T., Sedlmeier, P., Städtler, C., & Huron, D. (2013). The psychological functions of music listening. *Frontiers in Psychology*, *4*, 511. https://doi.org/10.3389/fpsyg.2013.00511

Scheer, M., Bülthoff, H. H., & Chuang, L. L. (2018). Auditory task irrelevance: A basis for inattentional deafness. *Human Factors*, *60*(3), 428–440. https://doi.org/10.1177%2F0018720818760919

Schröger, E., Giard, M. H., & Wolff, C. (2000). Auditory distraction: Event-related potential and behavioral indices. *Clinical Neurophysiology*, *111*(8), 1450–1460. https://doi.org/10.1016/S1388-2457(00)00337-0

Seli, P., Risko, E. F., Smilek, D., & Schacter, D. L. (2016). Mind-wandering with and without intention. *Trends in Cognitive Sciences*, *20*(8), 605–617. https://doi.org/10.1016/j.tics.2016.05.010

Stewart, H. J., Amitay, S., & Alain, C. (2017). Neural correlates of distraction and conflict resolution for nonverbal auditory events. *Scientific Reports*, *7*(1), 1–11. https://doi.org/10.1038/s41598-017-00811-7

Waltham-Smith, N. (2017). *Music and belonging between revolution and restoration*. Oxford: Oxford University Press.

Wegner, D. (1997). Why the mind wanders. In J. Cohen & J. Schooler (Eds.), *Scientific approaches to consciousness* (pp. 295–315). Mahwah, NJ: Erlbaum.

14 Musical Daydreaming and Kinds of Consciousness

Ruth Herbert

Here are excerpts from two accounts of music listening, separated by almost a century. The first is from what's perhaps one of the earliest empirical studies of music listening experiences, dating from the early 20th century. The second is drawn from my own research on the phenomenology of everyday music listening.

> [O]ne is not under ordinary conditions; one is not only day-dreaming, but day-dreaming to music . . . visions of superhuman creatures, of Elysian loveliness and cosmic grandeur; visions also of longings fulfilled in a vague future or a past no less imaginary and interesting. Such things are wafted into mind when . . . "under music."
>
> (Lee, 1933, pp. 278–279)

> You know just before you properly wake up in the morning . . . you're not asleep but you're not really back in the land of the living? Sometimes I can get like that with music. It's taking you away somewhere isn't it? [David, 51]
>
> (Herbert, 2011, p. 64)

Lee highlights content (fantastical visual mental imagery), clearly modulated by musical characteristics. By contrast, David describes a process where awareness of musical characteristics recedes; music affords a platform for consciousness transformation—specifically, to a stage between sleep and wakefulness (termed the hypnopompic state in sleep research). Both accounts emphasize inwardly directed attentional awareness and both could be conceptualized as examples of *mind-wandering/daydreaming* (Konishi, this volume), broadly understood as phenomena "characterized by spontaneous emergence of internally oriented images and thoughts largely unrelated to the present, external sensory environment" (Taruffi & Küssner, 2019, p. 64). But do *all* listening episodes involving internally oriented images involve a decoupling of attention from external surroundings?

This chapter examines first-hand reports of subjective experiences of listening to music I've accumulated over more than a decade, from several UK-based research projects centring on the phenomenology of everyday music listening.[1] I begin by considering definitions of music-evoked mind-wandering and musical daydreaming, and explain my adoption of the latter term, before exploring

DOI: 10.4324/9780429330070-18

emergent characteristics of musical daydreaming episodes apparent in verbal and written reports of subjective music-listening experiences, including sequential versus fragmented experience; absorption ("an effortless, non-volitional quality of deep involvement," as opposed to effortful attentional engagement; Jamieson, 2005, p. 120; see also Vroegh, this volume), and detachment (cutting off from self, task or environment; Herbert, 2011); repetitive daydreams; perceiver perspective (e.g., first or third person); multimodal experience; age-related differences in content and structure of experiences. I link examples to conceptualizations of kinds of consciousness (here understood as qualitatively different forms of awareness, including self-consciousness [awareness mediated by associations, memories and awareness of self], and a present-centred awareness [centred on sensory qualities of stimuli such as visual, auditory, etc.]) to phenomenological properties of non-musical mind-wandering/daydreaming identified in cognitive psychological literature (see Stawarczyk, 2018, for an overview).

Defining the Territory: Musical Daydreaming

Definitions of mind-wandering and daydreaming in music research centre on an internally directed attentional focus are illustrated as follows:

- Mind-wandering is "a form of self-generated thought which involves overcoming the constraints of the 'here and now' by immersing in one's own stream of consciousness" (Taruffi et al., 2017, p. 1).
- Daydreaming and mind-wandering "occur when attention is directed to internally generated thoughts and drifts away from the actual sensory environment" (Martarelli et al., 2016, p. 28).

Internally generated mental imagery is most often visual and auditory (visual and auditory imagery are the key focus of this chapter), but may be multimodal, including gustatory, olfactory, and tactile percepts (Taruffi & Küssner, 2019, p. 62). There is no consensus regarding terminology relating to classification of mental experiences (for a review relating to music, see Taruffi & Küssner, 2019). However, many researchers support the notion that the terms "mind-wandering" and "daydreaming" are interchangeable and refer to the same broad construct.

From a conceptual perspective, I adopt the term "daydreaming" in recognition of Jerome Singer's seminal reframing of daydreaming as a normative process, possessing creative and restorative potential (McMillan et al., 2013). Real-world studies of musical daydreaming are inclusive of episodes attached to broad ranges of tasks/activities, such as travelling, exercising, and relaxing/just "being." Music and mentation can intertwine throughout, or music may provide an induction to daydreaming, then fade from awareness.

Musical daydreams are frequently multimodal: they may feature a simultaneously distributed attentional focus (encompassing experiential blending of internal and external phenomena) or be marked by fluctuation between internal multimodal mentation (i.e., including visual imagery, verbal and non-verbal thought) *plus* awareness of external sensory environment. The multimodal nature of musical daydreaming accords with an ecological model of perception where experience is systemic—the sum of the interaction among individuals, stimulus attributes, and environment (Clarke, 2005).

Multimodal Listening and Kinds of Consciousness

Musical daydreams reflect a heteronomous, multimodal model of listening, sometimes called "hearing as": subjective experience is informed by extra-musical factors, including socio-cultural sources and autobiographical memories (Jakubowski, this volume) and aspects of external environment. Visual imagery (spontaneous or volitional) constitutes one potential modality (Taruffi & Küssner, this volume), alongside verbal and non-verbal mentation and multisensory perception. Heteronomous listening is frequently distinguished from autonomous listening, where music is sole focus of attention (deriving from Western Classical tradition, informed by notions of music as absolute, possessing immanent, and intramusical meaning via its structural properties; Scruton, 1997). Empirical evidence suggests that division between these modes is artificial; they may operate in tandem or fluctuate (Clarke, 2005; Herbert & Dibben, 2018). This was acknowledged by Vernon Lee in her pioneering qualitative listening study, earlier referred to:

> moments of concentrated and active attention to the musical shapes are like islands continually washed over by a shallow tide of other thoughts: memories, associations, suggestions, visual images and emotional states . . . they coalesce, forming a homogeneous and special contemplative condition. . . . Musical phrases, non-musical images and emotions are all welded into the same musical day-dream.
>
> (Lee, 1933, p. 32)

Lee highlights the processual nature of heteronomous listening episodes, with distributed attentional focus, fluctuating between awareness of musical attributes and extra-musical associations. Such episodes involve shifts of consciousness away from what individuals perceive as base-line "norm" towards a "special contemplative condition," the sum of interweaving webs of experiential phenomena. Multimodal music-listening may afford subtle shifts of consciousness precisely, *because* it simultaneously mobilizes and synthesizes different modes of experience.

I've suggested elsewhere (Herbert, 2011) that such interactions between music and consciousness may be usefully conceptualized as a form of musical trancing, "characterised by decreased orientation to consensual reality, decreased critical faculty, selective internal or external focus, together with changed sensory awareness" (p. 5). Trancing may be marked out by the presence of two psychological processes: absorption (non-volitional, effortless involvement) and dissociation (detachment). In addition, such listening episodes can be framed in terms of kinds of consciousness, from phenomenal consciousness, marked by "direct" sensory awareness, engaged with raw, ineffable qualities of subjective experience (qualia), to core (present focused) and extended (autobiographical) consciousness, informed by memory (Damasio, 1999). Conceptualizations of consciousness are—to greater or lesser degrees—theoretical constructions, but emphasize aspects of subjectivity highly relevant to music and mental imagery research, including the role of unconscious perception (non-verbal, preconscious processing) and the processual nature of unfolding, lived experiences of music.

Phenomenological Features Identified in Extant Literature

Temporal orientation and goal-relatedness have been key phenomenological foci of cognitive psychological research on mind-wandering/daydreaming. Other identified phenomenological features have received significantly less research attention, for example, frequency of sequential versus fragmented experiences, perceiver perspective (first- or third-person view, about self or "other"), repetitive daydreams, and relationship to context (Stawarczyk, 2018, pp. 207–208). Research on music-evoked mind-wandering (most of which centres on visual imagery; Taruffi & Küssner, this volume) has tended to focus on particular characteristics such as frequency and type of visual imagery (Küssner & Eerola, 2019), relationship between visual imagery and musical attributes (such as tempo and valence) plus neural correlates (e.g., the Default Mode Network; Taruffi et al., 2017). Far fewer studies have addressed phenomenological content of music-evoked mind-wandering and musical daydreams in terms of overall (holistic) subjective "feel" of lived experience or considered (in accord with systemic [ecological] understandings of experience) the potential impact of mediating factors such as age, context, and personality characteristics. I next discuss five excerpts from first-hand reports of subjective experiences with and of music to illustrate key phenomenological characteristics, relating them to kinds of consciousness.[2]

Musical Daydreams in Daily Life

Sequential Versus Fragmented Experience[3]

> On train listen to Shostakovich *Leningrad Symphony*. Always loved this for describing war horror—very filmy to me. Really feel hate and pain inside the music. Stare out of window, book unread, but probably not relating the views to music much. Internal mental images that I was getting were of horror of war from news footage. Lots of slow motion for some reason. Lots of thoughts and pictures about death and destruction mostly, but include frequent images of Shostakovich's face with square framed bakelite glasses and suit and tie and thinking how his appearance and the music seemed so opposite. [Max, 46]
>
> (Herbert, 2011, p. 70)

This episode possesses both narrative/sequential and fragmented qualities (apparent in dream-like tangential images/thoughts concerning Shostakovich's appearance). Max is a film recording mixer and professional musician, factors that serve to mediate subjective experience. In interview, he stated "if you have an emotional response to music *it will lead to images, whether you like it or not*" (2011, p. 70). His developed, habitual mode of response to music is listening visually, inward focus on imaginative involvement—even in eyes-open contexts. The influence of film is evident in his reference to slow motion sequences—a technique frequently used to heighten emotional involvement at dramatic/disturbing points in films. Cross-comparison of Max's reports revealed that self-chosen, familiar music appeared crucial to triggering involvement.[4] In terms of kinds of consciousness, the episode involves an extended awareness drawing on previous associations and memories. Diminishing connection from surroundings, and spontaneous increase in mental imagery, indicates increasing absorption (Vroegh, this volume). Reports from my real-world studies of musical subjectivity (Herbert, 2011) highlight travel as a key context for musical daydreaming.[5] Functionally, rather than relating to future planning or enhancement of creativity, such episodes appear self-regulatory—zoning out providing respite from mundane concerns.

"Hearing-as": Music as Virtual Person

> "I am sitting still for this extract- I feel that it wants me to listen. The music symbolises a nasty shock which has hurt someone very much. It sounds like an innocent being has been hurt by something or someone dreadful. [Female, 14]"
>
> (Herbert & Dibben, 2018, p. 384)

This quote is from a female student with high-level musical training (post-diploma) taking part in a lab-based listening study, where participants aged 10–18 were asked for free text responses to 20 extracts of experimenter-selected music.

The student is reacting to the atonal opening of Webern's *String Quartet op. 5 number 2*. Her description suggests that she experiences this as projecting person-like, frightening presence, possessing both intent ("it wants me to listen") and agency ("which has hurt someone"). Levinson (1997) proposes that hearing music this way correlates with empathic capacity; there is some evidence that empathic individuals experience higher levels of neural activity in the primary visual cortex (Taruffi et al., 2021). Musical attributes clearly connect to meaning-making (dissonance specifying fear/shock), demonstrating extended conscious awareness, informed by prior associations. Cross-comparison of reports indicates participants consistently associated dissonance with fear/shock, suggesting influence of musical codes encountered in film/TV. One main advantage (and purpose) of this lab-based study was that it facilitated assessment of shared levels of cultural understanding and meaning-making in relation to music across a sizeable sample of individuals, with different levels of exposure to music/musical training and socio-demographic backgrounds (N = 90). Intercultural consensus regarding types of visual imagery was evident across 90% of extracts (Herbert & Dibben, 2018, p. 385). However, the experimental setting and brief (up to 30 seconds) experimenter-selected extracts limited tapping of phenomenological detail of music-evoked imagery episodes. Findings should be considered alongside first-hand accounts of experiences of music in naturalistic contexts.

Age and Autobiographical Fantasy

> In one piece—something by "Basshunter." . . . I see myself in some random road . . . floating in the air, moving stuff with my mind . . . sort of controlling the weather. . . . And as the music progresses [it's] sort of a bit like a sci-fi movie. The actions and the drama gets more intense as the music gets to its climax . . . Basshunter's very techno-modern and it is easier to access it in the techno-modern music because it is very . . . sort of bass dominated so you can get very strong feelings from it and . . . if the volume's at a certain pitch I find it a lot easier to access that path. I can't do it with classical music. Just really big sounds. I can get really into my imagination. I have got a bit of an over active imagination!

Q. This is regular this accessing? . . .
R. Yes. Daily. Sort of alternate world sort of thing . . . because I don't really like *my* world a lot.
Q. And is access only through music?
R. Only through music.

[John, 17] (Herbert, 2019, p. 242)

Musical daydreaming is frequently equated with spontaneous, non-volitional thought, yet may equally be volitional (Christoff et al., 2016; Taruffi & Küssner, 2019). John, an "A" level music student planning a class music teaching career, engages in simultaneously absorbed and dissociative music-evoked mind-wandering daily, a practice identified in non-musical daydreaming literature, as particularly likely when episodes are autobiographical (Jakubowski, this volume), relating to personal goals (Stawarczyk, 2018) in addition to being prominent for young people aged 17–24 (Giambra, 2000).

John's habitual use of music to trigger repetitive imaginative autobiographical fantasies dissociates him from mundane concerns, allowing him to explore a powerful, autonomous identity. Musical qualities (musical attributes and extra-musical associations) tightly link to mental imagery. Loud volume and bass-dominated features of techno style specify power, plus associations with sci-fi films. Musical structure (building to climax) supports filmic, generally linear (albeit dream-like) visual narrative, in terms of perceiver perspective seen from third person viewpoint. As with Max, inwardly directed attentional focus, diminished awareness of surroundings, and upsurge of mental imagery point to extended conscious awareness. John clearly perceives this as a shift from his ordinary state of consciousness, an "alternate world," accessible only via music.

Age, Visual Imagery, Musical Attributes, and Musical Codes

> "I have a big habit of having daydreams, and when I'm listening to music . . . all the daydreams seem to come out. . . . When I was listening to 'Astor Piazzolla'[6] in the car there was this rather creepy track and I imagined there was someone being murdered—a small child actually, and there was this evil killer who we don't know of—no-one's ever seen their face as it's hidden under a black hood. And they're the one killing all these children. And it leaves. And it sees a little poor baby. It's had some trouble with being a child in its previous life and it thinks that all children are horrible due to what's happened to it. It looks at the child and it starts to feel sorry for the child. Then it forgets, throws it into the river and starts murdering a whole load of other kids . . . that's one of the stories. I can't really remember all of the stories I dream about. It's usually quite dramatic. [Lily, 11]"
>
> (Herbert, 2017, p. 154)

Reports from my study of young people's music-listening experiences (Herbert, 2019) suggest age impacts upon experience. From mid-adolescence, musical daydreaming with self-chosen music seems increasingly "about" the individual (i.e., autobiographical), involving imagined actual, idealized self, personal goals, and aspirations. By contrast, experiences from prepubescence through early adolescence (ca. aged 10–13 years) more likely feature fictional scenarios and characters, a sequential, narrative quality—a mode of engagement I term "storying" to music, as evident in Lily's description mentioned earlier. Similar to John, music affords an inductive platform to consciousness transformation, characterized by absorbed, imaginative involvement, informed by extended conscious awareness.

There is a relationship between visual imagery and structural properties of music (here an extended work in *nuevo tango* style, incorporating jazz and classical genre elements). By 11 years old, Lily (a formally trained pianist) is familiar with various musical codes (from a semiotic perspective, these would be termed "signifiers") through enculturation. It is not individual musical attributes—tempo, mode, and valence—prompting imaginative involvement but combined attributes heard holistically (semiotician Philip Tagg terms this a "museme stack"; Tagg & Clarida, 2003, p. 94).[7] Crucially, the piece features sound attributes closely prescribing sonic, kinetic, and tactile qualities of extra-musical sources (Tagg defines these as "anaphones," customizing the word "analogy"; Tagg & Clarida, 2003, p. 99). These include sinister bass drum ostinati, perhaps signifying hitting/thumping, plus rapid upward violin glissandi (reminiscent of Bernard Herrmann's use of strings in the famous shower scene from *Psycho*, 1960) suggesting rapid knife movements.

Both John and Lily's accounts show that music listening affords access to fantasy worlds, but there is a difference in attitude to musical daydreaming. John's account indicates awareness of *function*; a deliberate, repeated use of music to escape to "an alternate" world. Although Lily is aware of the term "daydream," such episodes occur spontaneously, as part of her everyday experience.

Multimodal Daydreaming, Distributed External/Internal Attentional Focus

> "It feels like I'm almost inside the video—not really participating in it, although I'm moving to it. . . . I can see all of the dancers, and all the things that I can remember from the video. As I can't remember it exactly, I also start to make bits of it up in my mind, but I can still see it all happening in front of me as well, they're all in different costumes . . . throughout all of this I can still look at the street—brick walls and green wheelie bins—things I wouldn't normally home in on—and not trip or bump into anything. [Alice, 16]"
>
> (Herbert, 2019, p. 240)

In contrast to the previous reports, Alice's description highlights distributed external and internal awareness. Alice has no formal musical training and does not play an instrument. She walks to a friend's house, listening to *Bad Girls*, by M.I.A., triggering memories of the video. The experience is multimodal (including sight, sound, mental visual imagery, and movement). The virtual, imagined world of the video (exotic visual images of Ouarzazarte, Morocco), sound of the track (Middle Eastern/Worldbeat), and external reality intersect; they are present in consciousness simultaneously, yet no experiential "blending" occurs, although heightening of external sensory awareness is evident in perception of small environmental details, normally unnoticed. This episode suggests interweaving of core (present centred) and extended kinds of consciousness. Alice had never reflected on her engagement with music prior to completing a listening diary. In interview, she revealed that doing so had highlighted a default approach to music listening dating back to childhood, unconsciously used to regulate mood/heighten energy levels.

Conclusions and Future Directions

Music-evoked daydreaming is commonly understood as involving inwardly directed focus on perceptual information accessed from memory or imagination and diminishing awareness of external non-musical phenomena. However, evidence from subjective reports of lived experiences of music demonstrates that sometimes attentional focus may fluctuate, inwards and outwards, or be simultaneously distributed between internal and external phenomena. A key characteristic is their multimodality, whether manifest in experiential blending of internal and external phenomena, or in fluctuation between internal mentation (including visual imagery, verbal and non-verbal thought) and awareness of external sensory environment. This systemic interaction between environment, perceiver, and affordances of music accords with ecological models of perception where subjective experience is modulated by a network of internal/external variables, including age, training, context, personality, and intention (daydreaming as deliberate or spontaneous).

I have highlighted age as an important mediator of musical experience, particularly during transition from early to late adolescence. Young people in the industrialized West appear particularly prone to "hearing-as," that is, heteronomous listening appears as default mode. Visual imagery is especially prevalent, with specific content influenced, in part, by gradual enculturation through repeated exposure to film and TV and computer games, where musical attributes are repeatedly paired with extra-musical stimuli (Herbert & Dibben, 2018; Schubert & McPherson, 2016). Specifically, cross-comparison of first-hand reports of musical daydreaming suggests a move towards autobiographical fantasy during mid-adolescence, although further research is needed to confirm this finding.

Another variable receiving limited research attention is the relationship between time of day and frequency/qualities of musical daydreams. I have argued that musical daydreaming affords subtle shifts of consciousness: modes of experience are synthesized in ways distinct from individuals' perceived baseline state of functioning (e.g., the replacement of one temporal frame [clock time] with another [a dream-like alternation between sequential and fragmented experience], or the alteration of perceiver perception to a dissociative third-person position). Musical daydreaming appears to privilege extended conscious awareness, albeit on occasion with core and extended kinds of consciousness interweaving. Findings from chronobiological literature indicate that conscious functioning is modulated by biological rhythms (both circadian [24 hour] and ultradian [recurrent short cycles throughout 24-hour periods]), with a small body of evidence supporting increase in mental imagery during recuperative "rest" phases of the ultradian cycle (Kleitman, 1982; Kokoszka, 2007). Chronobiological understanding of musical experience remains speculative (Bailes, 2019), but the hypothesis that musical daydreaming may function as a self-regulatory process affording respite from the vicissitudes of daily life—a space for simply "being"—merits further exploration.

Notes

1 My study of everyday music listening predominately focused on subjective experiences in real-world contexts, tapped via diary and interview methods. Purposive sampling was the main participant identification method, with age, level of involvement, and formal/informal training of the main criteria. Participant quotes appear in previous publications but analysis and discussion (utilising the concept of musical daydreaming) in this chapter are new (interpretative phenomenological analysis [IPA] was employed to reveal key themes).
2 Twenty-five reports were analysed utilising IPA. According to IPA aims, frequency *and* phenomenological richness were factors in the selection process. The five excerpts were selected as best representing emergent themes identified across the reports.
3 Subheadings are intended to highlight single themes that are particularly integral to each experience, not to represent all emergent thematic characteristics.
4 In an empirical study of narrative experiences of music, individuals were shown to be six times more likely to generate visual imagery when hearing familiar music (Margulis, 2017).
5 According to the findings of Jakubowski and Ghosh (2021) that music-evoked autobiographical memories most often occurred while travelling.
6 "Sextet," from *Luna* (1992) by Astor Piazzolla.
7 Tagg drew on Seeger's (1960) term *museme* (minimal unit of musical meaning), understood as equivalent to morpheme (minimal unit of linguistic meaning).

References

Bailes, F. (2019). Musical imagery and the temporality of consciousness. In R. Herbert, D. Clarke, & E. Clarke (Eds.), *Music and consciousness 2: Worlds, practices, modalities* (pp. 254–270). New York: Oxford University Press. https://doi.org/10.1093/oso/9780198804352.003.0016

Christoff, K., Irving, Z. C., Fox, K. C., Spreng, R. N., & Andrews-Hanna, J. R. (2016). Mind-wandering as spontaneous thought: A dynamic framework. *Nature Reviews Neuroscience*, *17*(11), 718–731. https://doi.org/10.1038/nrn.2016.113

Clarke, E. F. (2005). *Ways of listening: An ecological approach to the perception of musical meaning*. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195151947.001.0001

Damasio, A. (1999). *The feeling of what happens: Body, emotion and the making of consciousness*. London: Heinmann.

Giambra, L. M. (2000). The temporal setting, emotions, and imagery of daydreams: Age changes and age difference from late adolescent to the old-old. *Imagination, Cognition and Personality*, *19*(4), 367–413. https://doi.org/10.2190/h0w2-1792-jwuy-ku35

Herbert, R. (2011). *Everyday music listening: Absorption, dissociation and trancing*. Abingdon: Routledge. https://doi.org/10.4324/9781315581354

Herbert, R. (2017). Everyday trancing and musical daydreams. In R. Finnegan (Ed.), *Entrancement: The consciousness of dreaming, music and the world* (pp. 149–164). Cardiff: University of Wales Press.

Herbert, R. (2019). Absorption and openness to experience: An everyday tale of traits, states, and consciousness change with music. In R. Herbert, D. Clarke, & E. Clarke (Eds.), *Music and consciousness 2: Worlds, practices, modalities* (pp. 233–253). New York: Oxford University Press. https://doi.org/10.1093/oso/9780198804352.003.0014

Herbert, R., & Dibben, N. (2018). Making sense of music: Meanings 10- to 18-year-olds attach to experimenter-selected musical materials. *Psychology of Music*, *46*(3), 375–391. https://doi.org/10.1177/0305735617713118

Jakubowski, K., & Ghosh, A. (2021). Music-evoked autobiographical memories in everyday life. *Psychology of Music*, *49*(3), 649–666. https://doi.org/10.1177/0305735619888803

Jamieson, G. (2005). The modified Tellegen Absorption Scale: A clearer window on the structure and meaning of absorption. *Australian Journal of Clinical and Experimental Hypnosis*, *33*(2), 119–139.

Kleitman, N. (1982). Basic rest-activity cycle-22 years later. *Sleep*, *5*(4), 311–317.

Kokoszka, A. (2007). *States of consciousness: Models for psychology and psychotherapy*. New York: Springer. https://doi.org/10.1007/978-0-387-32758-7

Küssner, M. B., & Eerola, T. (2019). The content and functions of vivid and soothing visual imagery during music listening: Findings from a survey study. *Psychomusicology: Music, Mind, and Brain*, *29*(2–3), 90–99. http://dx.doi.org/10.1037/pmu0000238

Lee, V. (1933). *Music and its lovers: An empirical study of emotional and imaginative responses to music*. New York: E.P. Dutton.

Levinson, J. (1997). Music and negative emotion. In J. Robinson (Ed.), *Music and meaning* (pp. 215–241). Ithaca: Columbia University Press. https://doi.org/10.7591/9781501729737-012

Margulis, E. (2017). An exploratory study of narrative experiences of music. *Music Perception*, *35*(2), 235–248. https://doi.org/10.1525/mp.2017.35.2.235

Martarelli, C. S., Mayer, B., & Mast, F. W. (2016). Daydreams and trait affect: The role of the listener's state of mind in the emotional response to music. *Consciousness and Cognition*, *46*, 27–35. https://doi.org/10.1016/j.concog.2016.09.014

McMillan, R. L., Kaufman, S. B., & Singer, J. L. (2013). Ode to positive constructive daydreaming. *Frontiers in Psychology*, *4*, 1–9. https://doi.org/10.3389/fpsyg.2013.00626

Schubert, E., & McPherson, G. E. (2016). The perception of emotion in music. In G. E. McPherson (Ed.), *The child as musician: A handbook of musical development* (2nd ed., pp. 221–243). Oxford: Oxford University Press.

Scruton, R. (1997). *The aesthetics of music*. Oxford: Clarendon Press.

Seeger, C. (1960). On the moods of a music-logic. *Journal of the American Musicological Society*, *13*(1–3), 224–261. https://doi.org/10.2307/830257

Stawarczyk, D. (2018). Phenomenological properties of mind-wandering and daydreaming: A historical overview and functional correlates. In K. Christoff & K. Fox (Eds.), *The Oxford handbook of spontaneous thought: Mind-wandering, creativity, and dreaming* (pp. 193–214). New York: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190464745.013.18

Tagg, P., & Clarida, B. (2003). *Ten little title tunes: Towards a musicology of the mass media*. New York: Mass Media Music Scholars' Press.

Taruffi, L., & Küssner, M. (2019). A review of music-evoked visual mental imagery: Conceptual issues, relation to emotion, and functional outcome. *Psychomusicology: Music, Mind and Brain*, *29*(2–3), 62–74. https://doi.org/10.1037/pmu0000226

Taruffi, L., Pehrs, C., Skouras, S., & Koelsch, S. (2017). Effects of sad and happy music on mind-wandering and the default mode network. *Scientific Reports*, *7*(1), 1–10. http://dx.doi.org/10.1038/s41598-017-14849-0

Taruffi, L., Skouras, S., Pehrs, C., & Koelsch, S. (2021). Trait empathy shapes neural responses toward sad music. *Cognitive, Affective, & Behavioral Neuroscience*, *21*(1), 231–241. https://doi.org/10.3758/s13415-020-00861-x

15 Sound-Colour Synaesthesia and Music-Induced Visual Mental Imagery

Two Sides of the Same Coin?

Mats B. Küssner and Konstantina Orlandatou

Imagine taking part in an experiment about music-induced visual mental imagery. Together with other participants, you are being presented with a musical excerpt featuring a trumpet solo and are asked to report on the images you've seen in your mind's eye. Chatting with other participants about the experience after the experiment, you realize that your own visual mental imagery was quite different from theirs. While they mention landscapes, people dancing, or a trumpeter on stage, you experienced only distinct shapes of dark blue throughout the excerpt. You are surprised, because you always see these coloured shapes when listening to trumpet sounds and you thought that other people perceive them similarly.

This fictive example shows how an individual with (undetected) sound-colour synaesthesia may respond when being presented with trumpet sounds. Synaesthesia is a condition in which a stimulus (the "inducer") not only activates the appropriate sense—such as when sound activates hearing—but also consistently and involuntarily stimulates sensory pathways of another modality, that is, the "concurrent" (Simner & Hubbard, 2013). There is some evidence supporting the hypothesis that synaesthesia is a neurological condition caused by specific connections between different brain areas or unusual activation of cerebral areas (Mattingley et al., 2006; Nunn et al., 2002; Paulesu et al., 1995; Rouw & Scholte, 2007). The prevalence of synaesthesia is notoriously difficult to measure (Johnson et al., 2013), because there is currently no standardized test that fits all types of synaesthesia—over 150 different types have been reported—nor a uniform definition (Ward, 2013). While Simner and colleagues (2006) found that synaesthesia occurs in 4.4% of the population, other estimates range from 0.001% to 15% (Johnson et al., 2013). Synaesthetic (co-)activation can occur in all kinds of senses (hearing, sight, taste, touch, smell, etc.) and in any combination. It can manifest itself between several modalities (i.e., intermodally), such as coloured-hearing (audio-to-visual), within one modality (i.e., intramodally), such as grapheme-colour (visual-to-visual), or may also be triggered by a semantic concept or idea. An intriguing example of semantic inducers can be found in swimming-style synaesthesia, where the elicitation of the concept of a swimming style (e.g., by presenting a photograph) evokes a corresponding colour (see Nikolić et al., 2011). In all cases, an inducer (e.g., a sound) triggers the experience of a concurrent (e.g., a colour). In our aforementioned example, the sound-colour synaesthete may experience very distinct

DOI: 10.4324/9780429330070-19

shapes of dark blue every time she hears trumpet sounds. Some would argue that this specific co-activation of the visual pathway triggered by sounds is inborn, while others have identified "unconscious" rules which seem to be shared among some synaesthetes and which can also be found in cross-modal pairings of the general population (for a discussion, see Simner, 2013). It remains to be seen, then, whether our sound-colour synaesthete's sensory correspondence would have been always there or is the result of learned associations.

The existence of sound-colour synaesthesia has implications for studies on music-induced visual mental imagery (VMI), that is, the experience of seeing images in one's mind's eye while listening to music (see Taruffi & Küssner, this volume). Are sound-colour synaesthetes able to experience voluntary VMI on top of their synaesthetic experience? Do (sound-colour) synaesthetes experience music-induced VMI similar to non-synaesthetes when a musical excerpt does not trigger a synaesthetic response? And more generally, to what extent do subjective experiences of sound-colour synaesthetes and non-synaesthetes differ when music-induced VMI is concerned? Although research on synaesthesia and mental imagery is historically intertwined, little is known about the extent to which sound-colour synaesthesia and music-induced VMI tap the same cognitive mechanisms and how synaesthetic experiences may alter music-induced VMI. Comparing these two phenomena, we will argue that it may sometimes be difficult—or even impossible—to draw a clear line between synaesthetic and non-synaesthetic responses in contexts of music-induced VMI.

How Synaesthesia and Mental Imagery Are Related: A Very Brief History

Divided into two highly specialized research areas nowadays, the study of mental imagery and synaesthesia had a common origin in the 19th century. Although Galton did not use the term *synaesthesia*, his investigation of visualized numerals as part of his research on mental imagery (Galton, 1880) is now regarded as one of the first accounts of grapheme-colour synaesthesia (Ramachandran & Hubbard, 2001). Investigations of sound-colour synaesthesia (or chromesthesia), of which music-colour synaesthesia can be regarded as a subcategory, even date back to the beginning of the 19th century, comprising self-reports of the timbre-colour and tone-colour synaesthete Georg Tobias Ludwig Sachs (Simner, 2013). Studies explicitly referring to the term *synaesthesia* first appeared at the end of the 19th century, and most literature on synaesthesia around that time was concerned with sound-colour synaesthesia or "coloured-hearing" (Jewanski, 2013; Jewanski et al., 2020). However, as the behaviourist school gained increasing influence during the first half of the 20th century, the study of the mind and individuals' introspective accounts became less relevant. Only after the cognitive revolution in the 1960s do mental phenomena such as mental imagery re-enter the arena of psychological research. Interestingly, a definition of mental imagery by Richardson from 1969 would still encompass synaesthesia as it is being defined today (taken from Craver-Lemley & Reeves, 2013, p. 189):

> *Mental imagery refers to (1) all those quasi-sensory or quasi-perceptual experiences of which (2) we are self-consciously aware, and which (3) exist for us in the absence of those stimulus conditions that are known to produce their genuine sensory or perceptual counterparts, and which (4) may be expected to have different consequences from their sensory or perceptual counterparts*. (pp. 2–3).

While mental imagery—and particularly mental rotation tasks—constituted important research programmes in the 1970s and the 1980s, it took until the end of the 1980s before synaesthesia became *en vogue* again. The early work by Richard Cytowic was instrumental to re-establish the study of synaesthesia in the cognitive sciences. A breakthrough was his book *Synesthesia: A Union of the Senses* from 1989, which, for the first time, attempted to explain different neurological aspects of synaesthesia, as well as describing the different types and criteria of synaesthesia (Cytowic, 1989). Since his publication, research on synaesthesia has blossomed, including many studies with a focus on patterns of brain activity while subjects experience (grapheme-colour) synaesthesia (Nunn et al., 2002; Paulesu et al., 1995; Ramachandran & Hubbard, 2001; Rouw & Scholte, 2010; Ward et al., 2006), but also the development of standardized tests. Still unresolved is synaesthesia's relationship to other mental phenomena such as mental imagery, in order to identify conceptual commonalities and differences (Craver-Lemley & Reeves, 2013). Here, we argue that—in musical contexts—there is a case for regarding mental imagery as (weak) synaesthesia.

Mental Imagery as (Weak) Synaesthesia

While most people are able to see images in their mind's eye (but see Zeman et al., 2015), estimates of the prevalence of synaesthesia vary greatly, and many forms of synaesthesia may not yet have been discovered. A considerable number of cases could go undetected, because experiencing synaesthesia does not necessarily interfere with activities of everyday life. Synaesthetic or synaesthetic-like experiences might thus be more common in the general population than previously thought, and a visual concurrent and visual mental imagery may be two phenomena on a shared continuum of cross-modal experiences (Simner, 2013).

Both synaesthesia and mental imagery appear to have differential effects on perception. Craver-Lemley and Reeves (2013) have argued that synaesthesia can both facilitate and interfere with perception, whereas mental imagery predominantly interferes with perceptual tasks, also known as the "Perky effect." While Perky (1910), in her original study, showed that colours projected onto a wall can be mistaken for visual mental images, ever since the cognitive revolution researchers have extended this landmark study to demonstrate that the performance in perceptual detection or discrimination tasks in a given modality can be attenuated by concurrent imagery tasks in the same modality (for a review, see Waller et al., 2012). Further seemingly distinctive features between synaesthesia and mental imagery are vividness and controllability. Although there is evidence that vividness of mental imagery is greater in synaesthetes compared to non-synaesthetes (Barnett & Newell, 2008), there is no evidence that synaesthetes' vividness of mental images and vividness of synaesthetic experiences differ. Regarding controllability, synaesthesia appears to be distinct from mental imagery. Once the inducer is present—either in the external world or imagined—synaesthetes cannot suppress the automatic concurrent experience, whereas individuals experiencing mental imagery are generally able to control and manipulate their mental images. However, as we will see later, although synaesthetes cannot control the onset of the concurrent, some synaesthetes are able to do the equivalent of zooming in and out of visual concurrents.

As these comparisons show, there may be a case for regarding mental imagery as a (weak) form of synaesthesia (Martino & Marks, 2001), although this view has recently been disputed by others (Deroy & Spence, 2013; Spence, 2011). Martino and Marks (2001) made a separation between strong and weak synaesthesia suggesting that strong synaesthesia refers to inborn congenital synaesthesia, and weak synaesthesia refers to cross-modal correspondences underlying mechanisms of language, learning, and information processing. There is ample evidence that correspondences between different modalities are based on correspondences between common sensory dimensions, such as brightness, density, or size (Hornbostel, 1931; Karwoski et al., 1942; Marks, 1974; Simpson et al., 1956). For example, pitch correlates with visual size (Evans & Treisman, 2010; Gallace & Spence, 2006) and visual elevation (Ben-Artzi & Marks, 1995; Bernstein & Edelstein, 1971; Küssner et al., 2014; Küssner & Leech-Wilkinson, 2014; Melara & O'Brien, 1987), whereas loudness and pitch correlate with brightness (Argelander, 1927; Collier & Hubbard, 1998). Marks (2013) advocates that the main characteristic of weak synaesthesia is the degree to which one modal quality can be perceived like another; in other words, how the quality of one modality resembles the quality of a different modality as a form of intermodal analogy (Behne, 1991). For instance, in music-related correspondences, dark colours are associated with sounds of a low frequency range and bright colours to sounds of a high frequency range (Orlandatou, 2015). In the remainder of the chapter, we focus on sound-colour synaesthesia, which appears to be distinct from other types of synaesthesia. It may represent one example where the underlying mechanisms and phenomenal experiences shared with music-induced VMI are so closely related that the latter could be regarded as a weak form of sound-colour synaesthesia.

Sound-Colour Synaesthesia Versus Music-Induced Visual Mental Imagery

Jacobs and colleagues (1981) proposed that coloured-hearing can be the result of auditory information leakage out of the sensory pathway and its distribution to brain areas processing visual information. Some studies (de Thornley Head, 2006; Orlandatou, 2015; Ward et al., 2006) have shown that sound-colour synaesthesia is perceptual, since characteristics of the stimulus (e.g., pitch) have a true effect on the synaesthetic sensation. Additionally, it is believed that these sensations are due to a high-level perception rather than early unisensory information. Using a variation of the McGurk effect, Bargary et al. (2009) asked synaesthetes to describe the colour induced by a man's voice in a video. The audio channel of the video was manipulated by dubbing over an incongruent audio channel. The man pronounced the word "meat" (audio-visual), the manipulated audio was the word "neat" and the concluding *viseme* (i.e., visual counterpart of phoneme) was "peat." The colours induced by the illusion were different from the colours induced by the single auditory components, suggesting that early sensory activation is not directly linked to synaesthesia, but synaesthetic sensations are evoked after multimodal integration has occurred. However, many researchers claim that (sound-colour) synaesthesia might also be conceptual, based on cross-modal correspondences. It has been argued that cross-modal correspondences between different modalities are based on correspondences between common sensory dimensions, such as brightness or density, which are similar regardless of whether or not someone has synaesthesia (Marks, 1974; Ward et al., 2006). Could sound-colour synaesthesia and music-induced VMI thus be two sides of the same coin?

Curwen (2018) argues that music-colour synaesthetes are not always picked up by standardized tests such as the Synaesthesia Battery (Eagleman et al., 2007), because not all music-colour synaesthetes have tone-colour synaesthesia or similar. As the battery is designed to identify types of music-colour synaesthesias that are elicited by individual tones, chords, or a change of timbre, those types of music-colour synaesthesias that are not elicited from such stimuli but require, for example, the context of entire musical pieces or include the phenomenon of shapes and spatial landscapes are likely to be overlooked by such tests. The "colour" in music-colour synaesthesia is thus a proxy for a much richer perceptual experience than simply colour. There is also evidence that sound-colour synaesthetes have some level of control over their visual concurrents. For instance, Craver-Lemley and Reeves (2013) report a study with a synaesthete who was able to zoom in on a visual concurrent elicited by a synthesized tone. Thus, some forms of synaesthesia such as coloured-hearing may not fit common definitions of synaesthesia (Ward, 2013). They cannot be reliably detected with the Synaesthesia Battery, their visual concurrents are more complex and prone to variations if nuances of the inducer are changed, and some synaesthetes even show the ability to control the visual concurrents, presenting a striking similarity to music-induced visual mental imagery. Notably, there are reported cases of music-induced visual mental imagery whose features are synaesthesia-like. For instance, some participants in Küssner and Eerola's (2019) study reported experiencing very consistent visual images in response to specific musical pieces. Finally, in both phenomena, semantic inducers play an important role. For instance, an album cover can act as a semantic inducer of VMI: it can trigger mental images in the same way as listening to tracks of this album does. Similarly in synaesthesia, there is evidence that musical notation (Ward et al., 2006) or the concept of a tone without actually hearing it (de Thornley Head, 2006) can elicit visual concurrents in sound-colour synaesthetes. Despite these similarities it should be pointed out that the phenomenal experience of sound-/music-colour synaesthesia and music-induced visual mental imagery can sometimes be distinct.

Individuals experiencing music-induced VMI often mentally represent images in motion, since music is dynamic and comprises a set of events distributed in time. The same happens during synaesthetic sensation where synaesthetes often report that colours and forms are dynamic and not static. With VMI, individuals can represent the world in their minds and are able to recall past experiences or picture themselves in the future. In contrast, synaesthetic experiences have traditionally been conceptualized as not containing any reference to the external world, landscapes, persons or episodic memories, although this view has recently been disputed (Curwen, 2020). Furthermore, synaesthetic experiences do not occur deliberately but rather automatically and involuntarily. In a comparative study (Orlandatou, 2014), synaesthetes and non-synaesthetes were asked (1) to match colours as a correspondence to a sound, and (2) to explain why they made the specific correspondence. Non-synaesthetes could justify in every case their correspondences by trying to build a bridge between colour and sound through associative thinking such as “the sound is very low and like an airplane in blue sky” or an emotional response to the sound such as “pleasant” or “unpleasant.” Synaesthetes could describe their visual experiences by giving very precise and unusual descriptions such as “the sound starts on the right upper corner of my vision with a pretty clear black-green and fades out to white” or “a heavy black line and a lighter part of small grey particles flying around,” but they were unable to explain *why* they perceive these colours and formations. Such differences, if unaccounted for, can significantly influence the outcome of empirical studies. In the last section, we briefly consider how, in experiments of music-induced visual mental imagery, systematic differences between synaesthetes and non-synaesthetes may arise with regard to emotion, memory, and creativity.

How Synaesthetes’ and Non-Synaesthetes’ Responses to Music-Induced Visual Mental Imagery May Differ

Emotion

Several studies have shown that VMI is involved in emotional responses during music listening (Day & Thompson, 2019; Küssner & Eerola, 2019; Vuoskoski & Eerola, 2015). While synaesthetes often report emotions as a reference to whether the associations are right or wrong, for example, whether a grapheme is congruent or incongruent (Cytowic, 1989; Orlandatou, 2014), emotional response (e.g., to music) may also act as a semantic inducer in synaesthesia (Mroczko-Wąsowicz & Nikolić, 2014). A study investigating the associations between music, colour, and emotions revealed no differences between music-colour synaesthetes and non-synaesthetes (Isbilen & Krumhansl, 2016). Thus, if participants are able to freely conjure up visual imagery in response to music, there is no reason to assume that emotional responses between synaesthetes and non-synaesthetes will differ systematically.

Memory

Memory plays a significant role in music-induced visual mental imagery (see Jakubowski, this volume). People often remember a piece of music that accompanied important or salient events of their lives and automatically see the images of these events in their mind's eye when presented with that particular piece of music. In synaesthesia, however, memory does not seem to play a role for how synaesthetic experiences are triggered and manifested (but see again Curwen, 2020), although early accounts of synaesthesia often tried to find such associations (see Craver-Lemley & Reeves, 2013). Still, it is important to note that synaesthetes are sometimes found to have better memory than non-synaesthetes, particularly in some kinds of visual memory. For instance, grapheme-colour synaesthetes outperform matched controls in recognizing or recollecting abstract figures and patterns (Rothen et al., 2012). It is thus possible that synaesthetes are more susceptible to forming associations between episodic events and music and that their music-induced VMI—although not being a synaesthetic experience *per se*—is more often related to episodic memories than that of non-synaesthetes.

Creativity

Vividness of VMI has been shown to be positively related with a dimension of creativity (practicality of objects; Palmiero et al., 2011), and visual creativity and VMI seem to be closely related (Palmiero et al., 2015). There is also evidence that synaesthetes are capable of generating unusual and rather irrelevant associations between diverse concepts (Ward et al., 2008). Many artists such as the composer Olivier Messiaen and visual artist David Hockney have been motivated and inspired by their synaesthetic states. Messiaen (1944) described his compositional approach and tools for his works in the treatise "La Technique de mon langage musical." Many of his compositions are based on a system he called "modes of limited transposition." These modes refer to seven musical scales which Messiaen associated with colours—an approach that gave him the variety needed to play with harmonies of colour and music. Hockney worked on a series of stage designs for the Metropolitan Opera of New York. For the operas *The Magic Flute* by Mozart and *Turandot* by Puccini, he painted the sets according to his synaesthetic sensations evoked when he listened to the music of each piece. Research by Mulvenna (2013) suggests that there are indeed many synaesthetes among artists but that synaesthetes are not more likely to produce more creative artworks than non-synaesthetes. Rather, synaesthetes show more creative thinking—which can be seen as distinct from creative production (i.e., the presentation of an idea in a secondary medium)—than non-synaesthetes, suggesting that there are neural mechanisms underlying both synaesthetic experiences and creative cognition. In experiments of music-induced VMI, both non-synaesthetes with a high score on a vividness of visual imagery scale as well as synaesthetes can therefore be expected to come up with creative, perhaps unusual VMI during music listening.

Conclusion

Mental imagery and synaesthesia are both situated in highly specialized fields of study, even though there exist some striking similarities between these two phenomena. We have argued that music-induced visual mental imagery could be regarded as a weak form of sound-colour synaesthesia because some, but not all, conceptual and experiential features overlap. Are sound-colour synaesthesia and music-induced VMI therefore two sides of the same coin? The answer depends partly on what kind of working definition of synaesthesia one adopts. What seems clear is that synaesthetes' responses in experiments on music-induced VMI might differ from those of non-synaesthetes. To echo Price's (2013) suggestion that researchers should check synaesthetes' mental imagery abilities in experimental studies to account for potential confounds, we propose that researchers investigating mental imagery—and more particularly, music-induced VMI—should assess whether their participants are synaesthetes, using a combination of standardized tests and customized questions depending on the particular research context. Future research shedding light on synaesthetes' responses in studies of music-induced VMI may provide new insights into the origins and interindividual differences of how people form VMI in response to music.

References

Argelander, A. (1927). *Das Farbenhören und der synästhetische Faktor der Wahrnehmung* [The hearing of colors and the synesthetic factor of perception]. Jena: Verlag Gustav Fischer.

Bargary, G., Barnett, K. J., Mitchell, K. J., & Newell, F. N. (2009). Colored-speech synaesthesia is triggered by multisensory, not unisensory, perception. *Psychological Science*, *20*(5), 529–533. https://doi.org/10.1111/j.1467-9280.2009.02338.x

Barnett, K. J., & Newell, F. N. (2008). Synaesthesia is associated with enhanced, self-rated visual imagery. *Consciousness and Cognition*, *17*(3), 1032–1039. https://doi.org/10.1016/j.concog.2007.05.011

Behne, K.-E. (1991). Am Rande der Musik: Synästhesien, Bilder, Farben, . . . [On the edge of music: Synesthesias, images, colors, . . .]. In K.-E. Behne, G. Kleinen, & H. de la Motte-Haber (Eds.), *Musikpsychologie: Jahrbuch der Deutschen Gesellschaft für Musikpsychologie* (Vol. 8, pp. 94–120). Wilhelmshaven: Florian Noetzel Verlag.

Ben-Artzi, E., & Marks, L. E. (1995). Visual-auditory interaction in speeded classification: Role of stimulus difference. *Perception & Psychophysics*, *57*(8), 1151–1162. https://doi.org/10.3758/BF03208371

Bernstein, I. H., & Edelstein, B. A. (1971). Effects of some variations in auditory input upon visual choice reaction time. *Journal of Experimental Psychology*, *87*(2), 241–247. https://doi.org/10.1037/h0030524

Collier, W. G., & Hubbard, T. L. (1998). Judgments of happiness, brightness, speed and tempo change of auditory stimuli varying in pitch and tempo. *Psychomusicology: Music, Mind & Brain*, *17*(1), 36–55. https://doi.org/10.1037/h0094060

Craver-Lemley, C., & Reeves, A. (2013). Is synesthesia a form of mental imagery? In S. Lacey & R. Lawson (Eds.), *Multisensory imagery* (pp. 185–206). New York: Springer. https://doi.org/10.1007/978-1-4614-5879-1_10

Curwen, C. (2018). Music-colour synaesthesia: Concept, context and qualia. *Consciousness and Cognition*, *61*, 94–106. https://doi.org/10.1016/j.concog.2018.04.005

Curwen, C. (2020). Music-colour synaesthesia: A sensorimotor account. *Musicae Scientiae*, 1029864920956295. https://doi.org/10.1177/1029864920956295

Cytowic, R. E. (1989). *Synesthesia: A union of the senses*. New York: Springer. https://doi.org/10.1007/978-1-4612-3542-2

Day, R. A., & Thompson, W. F. (2019). Measuring the onset of experiences of emotion and imagery in response to music. *Psychomusicology: Music, Mind, and Brain*, *29*(2–3), 75–89. https://doi.org/10.1037/pmu0000220

Deroy, O., & Spence, C. (2013). Why we are not all synesthetes (not even weakly so). *Psychonomic Bulletin & Review*, *20*(4), 643–664. https://doi.org/10.3758/s13423-013-0387-2

de Thornley Head, P. (2006). Synaesthesia: Pitch-colour isomorphism in RGB-space? *Cortex*, *42*(2), 164–174. https://doi.org/10.1016/S0010-9452(08)70341-1

Eagleman, D. M., Kagan, A. D., Nelson, S. S., Sagaram, D., & Sarma, A. K. (2007). A standardized test battery for the study of synesthesia. *Journal of Neuroscience Methods*, *159*(1), 139–145. https://doi.org/10.1016/j.jneumeth.2006.07.012

Evans, K. K., & Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features. *Journal of Vision*, *10*(1), 1–12. https://doi.org/10.1167/10.1.6

Gallace, A., & Spence, C. (2006). Multisensory synesthetic interactions in the speeded classification of visual size. *Attention, Perception, & Psychophysics*, *68*(7), 1191–1203. https://doi.org/10.3758/BF03193720

Galton, F. (1880). Visualised numerals. *Nature*, *21*(533), 252–256. https://doi.org/10.1038/021252a0

Hornbostel, E. M. v. (1931). Über Geruchshelligkeit [About odor brightness]. *Pflüger's Archiv für die gesamte Physiologie des Menschen und der Tiere*, *227*(1), 517–538. https://doi.org/10.1007/BF01755351

Isbilen, E. S., & Krumhansl, C. L. (2016). The color of music: Emotion-mediated associations to Bach's well-tempered clavier. *Psychomusicology: Music, Mind, and Brain*, *26*(2), 149–161. https://doi.org/10.1037/pmu0000147

Jacobs, L., Karpik, A., Bozian, D., & Gøthgen, S. (1981). Auditory-visual synesthesia sound-induced photisms. *Archives of Neurology*, *38*(4), 211–216. https://doi.org/10.1001/archneur.1981.00510040037005

Jewanski, J. (2013). *Synesthesia in the nineteenth century*. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199603329.013.0019

Jewanski, J., Simner, J., Day, S. A., Rothen, N., & Ward, J. (2020). The evolution of the concept of synesthesia in the nineteenth century as revealed through the history of its name. *Journal of the History of the Neurosciences*, *29*(3), 259–285. https://doi.org/10.1080/0964704X.2019.1675422

Johnson, D., Allison, C., & Baron-Cohen, S. (2013). *The prevalence of synesthesia*. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199603329.013.0001

Karwoski, T. F., Odbert, H. S., & Osgood, C. E. (1942). Studies in synesthetic thinking: II. The role of form in visual responses to music. *The Journal of General Psychology*, *26*(2), 199–222. https://doi.org/10.1080/00221309.1942.10545166

Küssner, M. B., & Eerola, T. (2019). The content and functions of vivid and soothing visual imagery during music listening: Findings from a survey study. *Psychomusicology: Music, Mind, and Brain*, *29*(2–3), 90–99. https://doi.org/10.1037/pmu0000238

Küssner, M. B., & Leech-Wilkinson, D. (2014). Investigating the influence of musical training on cross-modal correspondences and sensorimotor skills in a real-time drawing paradigm. *Psychology of Music*, *42*(3), 448–469. https://doi.org/10.1177/0305735613482022

Küssner, M. B., Tidhar, D., Prior, H. M., & Leech-Wilkinson, D. (2014). Musicians are more consistent: Gestural cross-modal mappings of pitch, loudness and tempo in real-time. *Frontiers in Psychology*, *5*(789). https://doi.org/10.3389/fpsyg.2014.00789

Marks, L. E. (1974). On associations of light and sound: The mediation of brightness, pitch, and loudness. *The American Journal of Psychology*, *87*(1/2), 173–188. https://doi.org/10.2307/1422011

Marks, L. E. (2013). *Weak synesthesia in perception and language*. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199603329.013.0038

Martino, G., & Marks, L. E. (2001). Synesthesia: Strong and weak. *Current Directions in Psychological Science*, *10*(2), 61–65. https://doi.org/10.1111%2F1467-8721.00116

Mattingley, J. B., Payne, J. M., & Rich, A. N. (2006). Attentional load attenuates synaesthetic priming effects in grapheme-colour synaesthesia. *Cortex*, *42*(2), 213–221. https://doi.org/10.1016/S0010-9452(08)70346-0

Melara, R. D., & O'Brien, T. P. (1987). Interaction between synesthetically corresponding dimensions. *Journal of Experimental Psychology: General*, *116*(4), 323–336. https://doi.org/10.1037/0096-3445.116.4.323

Messiaen, O. (1944). *La Technique de mon langage musical* [The technique of my musical language]. Paris: Alphonse Leduc Éditions Musicales.

Mroczko-Wąsowicz, A., & Nikolić, D. (2014). Semantic mechanisms may be responsible for developing synesthesia. *Frontiers in Human Neuroscience*, *8*, 509. https://doi.org/10.3389/fnhum.2014.00509

Mulvenna, C. (2013). Synesthesia and creativity. In J. Simner & E. Hubbard (Eds.), *Oxford handbook of synesthesia* (pp. 607–630). Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199603329.013.0031

Nikolić, D., Jürgens, U. M., Rothen, N., Meier, B., & Mroczko, A. (2011). Swimming-style synesthesia. *Cortex*, *47*(7), 874–879. https://doi.org/10.1016/j.cortex.2011.02.008

Nunn, J. A., Gregory, L. J., Brammer, M., Williams, S. C. R., Parslow, D. M., Morgan, M. J., . . . Gray, J. A. (2002). Functional magnetic resonance imaging of synesthesia: Activation of V4/V8 by spoken words. *Nature Neuroscience*, *5*(4), 371–375. https://doi.org/10.1038/nn818

Orlandatou, K. (2014). *Synaesthetic and intermodal audio-visual perception: An experimental research* [Doctoral Dissertation, Staats-und Universitätsbibliothek Hamburg Carl von Ossietzky]. ediss.sub.hamburg. www.sub.uni-hamburg.de/startseite.html

Orlandatou, K. (2015). Sound characteristics which affect attributes of the synaesthetic visual experience. *Musicae Scientiae*, *19*(4), 389–401. https://doi.org/10.1177/1029864915593334

Palmiero, M., Cardi, V., & Belardinelli, M. O. (2011). The role of vividness of visual mental imagery on different dimensions of creativity. *Creativity Research Journal*, *23*(4), 372–375. https://doi.org/10.1080/10400419.2011.621857

Palmiero, M., Nori, R., Aloisi, V., Ferrara, M., & Piccardi, L. (2015). Domain-specificity of creativity: A study on the relationship between visual creativity and visual mental imagery. *Frontiers in Psychology*, *6*, 1870. https://doi.org/10.3389/fpsyg.2015.01870

Paulesu, E., Harrison, J., Baron-Cohen, S., Watson, J. D. G., Goldstein, L., Heather, J., . . . Frith, C. D. (1995). The physiology of coloured hearing: A PET activation study of colour-word synaesthesia. *Brain*, *118*(3), 661–676. https://doi.org/10.1093/brain/118.3.661

Perky, C. W. (1910). An experimental study of imagination. *The American Journal of Psychology*, *21*(3), 422–452. https://doi.org/10.2307/1413350

Price, M. C. (2013). Synesthesia, imagery, and performance. In J. Simner & E. Hubbard (Eds.), *Oxford handbook of synesthesia* (pp. 728–760). Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199603329.013.0037

Ramachandran, V. S., & Hubbard, E. M. (2001). Psychophysical investigations into the neural basis of synaesthesia. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *268*(1470), 979–983. https://doi.org/10.1098/rspb.2000.1576

Rothen, N., Meier, B., & Ward, J. (2012). Enhanced memory ability: Insights from synaesthesia. *Neuroscience & Biobehavioral Reviews*, *36*(8), 1952–1963. https://doi.org/10.1016/j.neubiorev.2012.05.004

Rouw, R., & Scholte, H. S. (2007). Increased structural connectivity in grapheme-color synesthesia. *Nature Neuroscience*, *10*(6), 792–797. https://doi.org/10.1038/nn1906

Rouw, R., & Scholte, H. S. (2010). Neural basis of individual differences in synesthetic experiences. *Journal of Neuroscience*, *30*(18), 6205–6213. https://doi.org/10.1523/JNEUROSCI.3444-09.2010

Simner, J. (2013). The 'rules' of synesthesia. In J. Simner & E. Hubbard (Eds.), *Oxford handbook of synesthesia* (pp. 149–164). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199603329.013.0008

Simner, J., & Hubbard, E. M. (Eds.). (2013). *Oxford handbook of synesthesia*. Oxford: Oxford University Press.

Simner, J., Mulvenna, C., Sagiv, N., Tsakanikos, E., Witherby, S. A., Fraser, C., . . . Ward, J. (2006). Synaesthesia: The prevalence of atypical cross-modal experiences. *Perception*, *35*(8), 1024–1033. https://doi.org/10.1068/p5469

Simpson, R. H., Quinn, M., & Ausubel, D. P. (1956). Synesthesia in children: Association of colors with pure tone frequencies. *The Journal of Genetic Psychology*, *89*(1), 95–103. https://doi.org/10.1080/00221325.1956.10532990

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, *73*(4), 971–995. https://doi.org/10.3758/s13414-010-0073-7

Vuoskoski, J. K., & Eerola, T. (2015). Extramusical information contributes to emotions induced by music. *Psychology of Music*, *43*(2), 262–274. https://doi.org/10.1177/0305735613502373

Waller, D., Schweitzer, J. R., Brunton, J. R., & Knudson, R. M. (2012). A century of imagery research: Reflections on Cheves Perky's contribution to our understanding of mental imagery. *The American Journal of Psychology*, *125*(3), 291–305. https://doi.org/10.5406/amerjpsyc.125.3.0291

Ward, J. (2013). Synesthesia: Where have we been? Where are we going? In J. Simner & E. Hubbard (Eds.), *Oxford handbook of synesthesia* (pp. 1022–1040). Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199603329.013.0049

Ward, J., Huckstep, B., & Tsakanikos, E. (2006). Sound-colour synaesthesia: To what extent does it use cross-modal mechanisms common to us all? *Cortex*, *42*(2), 264–280. https://doi.org/10.1016/S0010-9452(08)70352-6

Ward, J., Moore, S., Thompson-Lake, D., Salih, S., & Beck, B. (2008). The aesthetic appeal of auditory-visual synaesthetic perceptions in people without synaesthesia. *Perception*, *37*(8), 1285–1296. https://doi.org/10.1068%2Fp5815

Ward, J., Tsakanikos, E., & Bray, A. (2006). Synaesthesia for reading and playing musical notes. *Neurocase*, *12*(1), 27–34. https://doi.org/10.1080/13554790500473672

Zeman, A., Dewar, M., & Della Sala, S. (2015). Lives without imagery: Congenital aphantasia. *Cortex*, *73*, 378–380. https://doi.org/10.1016/j.cortex.2015.05.019

16 Music-Evoked Imagery in an Absorbed State of Mind

A Bayesian Network Approach

Thijs Vroegh

Music has long been linked to trance-like states of consciousness, including absorption (Becker, 2004; Fachner, this volume; Herbert, 2011, this volume). Building on the work by Tellegen (1981, 1992), *absorption* can be defined as a discrete-like state that is characterized especially (albeit not exclusively!) by a heightened and effortless form of attention together with a loss of a generalized reality orientation, exhibited by an altered awareness of time, body, surroundings, and a forgetting of oneself (see Butler, 2004; Herbert, 2011; Vroegh, 2019). The term "discrete-like" is used here to emphasize that absorption—in contrast to broader constructs such as *involvement* or *engagement*—refers to a specific subclass that involves high-intensity musical experiences only. At the same time, this definition avoids absorption to be an all-or-nothing-state that is categorically distinguishable from other states of consciousness (cf. Huskey et al., 2018). Instead, this chapter follows the view that consciousness states are continuous, blending into one another (Herbert, 2011). Demarcating the subclass of absorption thus necessarily is a matter of imprecise estimation rather than being marked by a clear observable on- and offset in attention and awareness (Vroegh, 2018).

The experiential dimensions that are most central in characterizing absorption (i.e., attention, altered experience, and loss of self-awareness) echo some of the basic phenomenal features of consciousness as identified by for instance Pekala (1991) and Rainville and Price (2003). Indeed, motivated to go beyond theoretical debates towards a more empirical-oriented approach, a fruitful line of thinking taken in psychological research is by construing states of consciousness in multidimensional terms (Farthing, 1992; Pekala, 1991; Studerus et al., 2010). From this perspective, consciousness states can be distinguished from other states in terms of *intensity* and the configurational patterns or *interrelations* between these basic experiential dimensions (Pekala, 1991; Tart, 1980). Ultimately, approaching absorption experiences from such a multidimensional take on consciousness has the advantage of properly dealing with, as Herbert (2019, pp. 233–234) calls it, "the holistic, subjective feel of individual experience" that can be described as "a network of physiological, emotional, cognitive and perceptual interactions."

Besides other dimensions such as internal dialogue, emotions, arousal, and memory, *visual imagery* is often regarded as another important experiential attribute of phenomenal consciousness (Pekala, 1991). Moreover, visual imagery is

DOI: 10.4324/9780429330070-20

considered to be an important part of the listening experience (Küssner & Eerola, 2019; Taruffi & Küssner, 2019, this volume),[1] and is often found to be prevalent in reports on absorption while listening to music (Gabrielsson & Wik, 2003; Herbert, 2011; Vroegh, 2019). Research on imagery typically lied emphasis on detailed descriptions of music experiences, highlighting both the diversity of imagery in terms of its contents and occurrences (Herbert, 2011) and that its function may vary depending on the circumstances of listening (Vuoskoski & Eerola, 2015). Moreover, in terms of interrelations with other aspects of the music experience, imagery was found to manifest itself differently depending on type of felt emotions during the absorbed engagement with music (Vroegh, 2019).

Although often linked to other experiential features such as emotions, thoughts, and episodic memory (Juslin & Västfjäll, 2008), it still remains unclear what the role is of imagery within the larger multidimensional framework of an absorbed state of mind. For this we need an experiential model that consists of multiple dimensions of subjective experience and that includes visual imagery. To this end, this chapter reports on a Bayesian network approach that can inform on how such a model may look like and how it can help to further examine imagery in the scenario of an absorbed listening state.

Bayesian Networks

A potential method from which to analyse the phenomenological structure of absorption paralleling the approach described earlier is *network analysis*. In line with the increasing recognition and application of a network approach to understand a variety of psychological phenomena and attitudes (Schmittmann et al., 2013), such a network perspective defines a state of mind as the emergent consequence of dynamically interacting and possible self-reinforcing experiential attributes. Conceptualized as a complex network system, consciousness thus is *mereological*—part(s) to whole (Borsboom & Cramer, 2013), with a state of consciousness defined as emerging from the specific interactions among given dimensions of consciousness rather than being the latent cause of its observable manifestations (cf. Agarwal & Karahanna, 2000).

An analysis technique useful in this context is Bayesian network analysis (Korb & Nicholson, 2010). A Bayesian network (BN) is a probabilistic graphical model that describes the joint probability distribution for a set of variables. The structure of the network is more generally referred to as a directed acyclic graph (DAG) and useful for representing causal relationships. This DAG comprises the conditional probability distributions of the variables of interest (*nodes*) and directed arrows (*edges*), encoding the statistical dependences and independences between them (see Figure 16.1). Each node can adopt different "states" (here in terms of three probability intervals *low, middle* or *high*) which can only be influenced by values of its parent nodes based on the structure of the graph. Given information of one or more variables that are part of the BN, the most probable values of other variables in the BN can be estimated. A DAG does not contain cycles, so that a node cannot effect itself through a feedback loop.

Methodological developments in network analysis also allow for computing Bayesian networks based on subjective, self-report data. Recently, a first study (Vroegh, 2021) applied a BN approach in the context of listening to a researcher-selected piece of music and investigating effects on the structure of phenomenal consciousness. Findings from this study underlined that imagery has an intermediary function linking together emotions, short- and long-term memory, and inward-directed attention. Importantly, some directional links involving imagery were more pronounced than others. For instance, whereas attentional direction clearly emerged as "causing" visual imagery, and positive affect and imagery both clearly affected "mixed affect," the link found between positive affect and imagery suggested a reciprocal effect. The remainder of this chapter will extend on these findings by reporting on a similar type of study with self-selected music.

The Scenario of an Absorbed Listening State

To empirically investigate how visual imagery manifests itself in a state of music-induced absorption, first the overall Bayesian network and the probability distributions is computed for the domain of music experience. Based on the resulting structure, we then focus in more detail on the scenario of being absorbed. We do this by providing *evidence* to the model by means of setting variables in known states (i.e., simulated intervention) and then, conditioned on this evidence, investigate its effects on the whole network by querying the posterior probability of imagery and other features in the network.

To this end, an online study was conducted in which one was requested to listen to a self-chosen, favourite piece of music of around 5 minutes. Directly afterwards, the participants (N=193) have to fill in the Phenomenology of Consciousness Inventory (PCI; Pekala, 1991). The PCI covers 12 dimensions that together represent a state of consciousness and among others includes attention, self-awareness, short-term memory, self-control, altered experience, positive and negative affect, visual imagery, internal dialogue, rationality, and arousal. In addition, several items in similar format (7-point bipolar differential scale) were included to tap into other dimensions of consciousness deemed relevant but which are not originally included in the PCI (i.e., sense of unity, reflective thoughts, mixed affect, and long-term memory).[2]

The structure of the network was "learned" from the available data by using the *hillclimbing* algorithm as part of the R package *bnlearn* (Scutari & Denis, 2015). Except for fixing the edge "attention → altered experience" where the arrow means direction of influence (Vroegh, 2018),[3] the learning process took place without any further prior-made assumptions about how the dimensions of consciousnesses should interrelate. Model-averaging was used to further increase the robustness of the model.[4] Final estimation of the probabilistic parameters was then performed with GeNie 2.3 (Druzdzel, 1999), which provides a suitable platform for further visual analysis. For a more detailed description of the different steps of the data-analytical procedure, the reader is referred to Vroegh (2021).

Figure 16.1 represents the resulting Bayesian network. It proposes several conditional interactions (i.e., directed arrows or edges) between the nodes matching

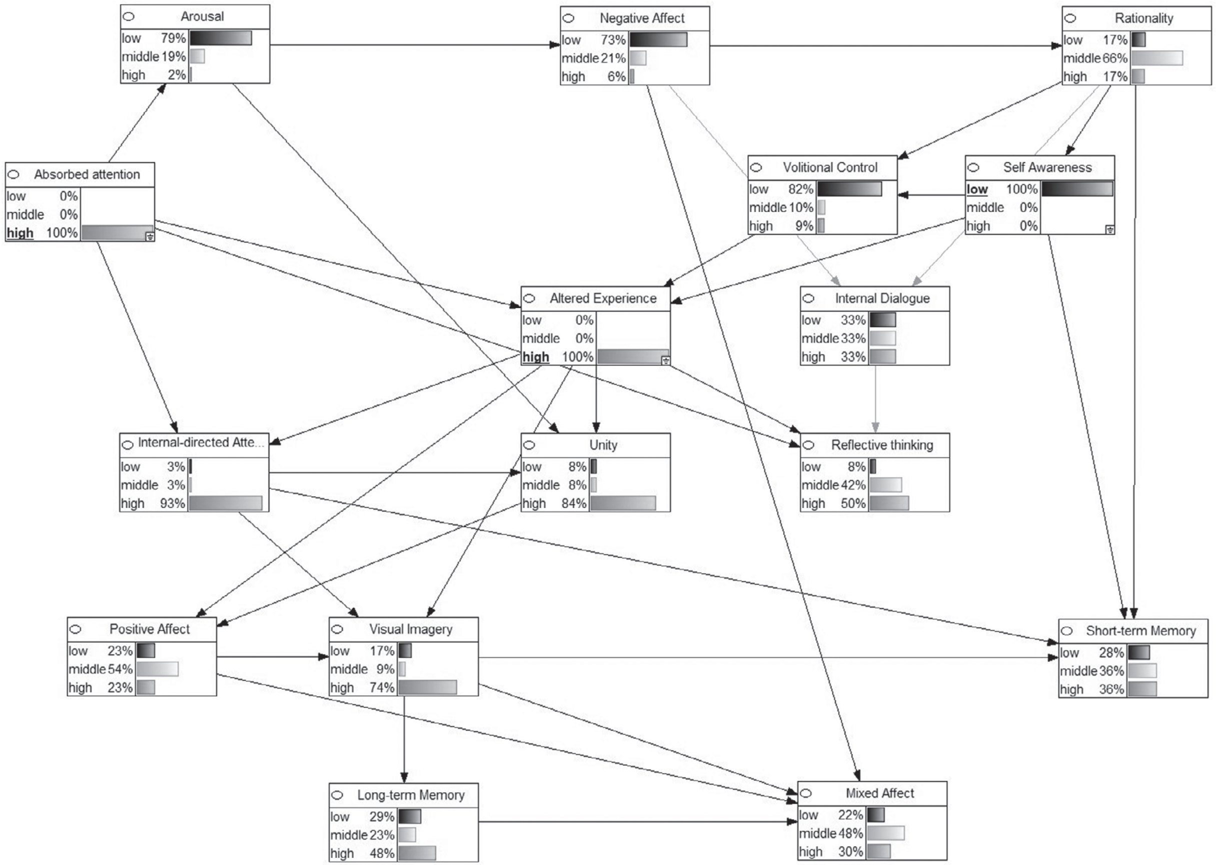


Figure 16.1 Bayesian network model of music absorption.

the dimensions of consciousness as formulated by Pekala (1991), Tart (1980) and Farthing (1992). Also, the probability distributions for each variable are estimated, with *low*, *middle*, and *high* representing the possible categories. In addition, following the definition of absorption outlined earlier, the figure depicts the specific "what-if" scenario of being absorbed by having set/instantiated absorbed attention and altered experience to *high* (at 100%; see boxes with "high" underlined), and self-awareness to *low* (at 100%). Correspondingly, the posterior probabilities (i.e., the probabilities *after* having provided this *evidence* of state absorption to the model) of the other variables in the network change. That is, *given* full attention, altered experience and being completely unaware of oneself (i.e., the three hallmark features of absorption), we can derive the probability of all the other included experiential attributes for this scenario. In this case, the probability of having visual mental imagery by participants who were completely absorbed while listening to a self-chosen piece of music reaches a high of 74%, based on the subjectively report by the music listener.

Looking further at the sparse structure of the network, altered experience emerges as a central "hub" within the network, whereas mixed affect, short-term memory, and reflective thoughts are identified as outcome attributes. Moreover, several links found by the learning algorithm immediately make intuitive sense, and are consistent with previous theoretical and empirical work. For instance, negative affect (which includes sadness) and rationality (i.e., logic reasoning) have repeatedly been related to each other (for an overview, see Pham, 2007). Likewise, the proposed dependency of mixed affect on positive- and negative affect in a v-structure (Scutari & Denis, 2015) is in line with conceptualizing mixed affect as a combination of basic emotions (Larsen & McGraw, 2011). The ability to remember exactly what happened while listening (i.e., short-term memory) is intuitively depending on the node of self-awareness. Also, the (possibly, causal) dependency of imagery on inward-directed attention makes intuitive sense (cf. Tellegen, 1992), just as the proposed directional association between imagery and altered experience. At the same time, established links that could be expected based on the literature such as "internal dialogue ↔ self-awareness," are surprisingly lacking (Morin, 2005). Sometimes, associations between nodes appear to be reversed (e.g., "positive affect → imagery"; cf. Juslin & Västfjäll, 2008), or else counterintuitively directed (e.g., it would be more logical to expect the edge "long-term memory → imagery"; Taruffi & Küssner, 2019).

The findings on attentional direction clearly suggests that an absorbed state goes together with an inward-oriented focus (a 93% probability of *high* in a state of absorption as compared with a 66% probability *before* providing the evidence of absorption to the model). In addition, the feeling of "merging with the music" as expressed by the unity node further suggests that the network appears to successfully represent a state of absorption (*high* was 84%). This means that the perceived feeling of merging—which can be interpreted as a useful index summarizing one's overall absorbed state of mind—takes place *within* the listener. Correspondingly, the probability of music listeners to report experiencing visual imagery during an intense state of absorption increases considerably (*high* increases from a prior probability of 40% to the posterior probability of 74%). These results are in line with Tellegen's (1981, 1992) conceptualization of absorption as often being reflected by an *inner focus* on reminiscences, images, and imaginings.

Also noteworthy is the finding that, although self-awareness was set to low to instantiate the model to a state of absorption, the short-term memory of the previous experience was not affected as much when compared to the increased sense of losing control (i.e., *low* of volitional control at 82%). This is consistent with research on strong experiences which are often marked by absorption but which are nonetheless often quite well remembered even when they happened a long time ago (Gabrielsson & Wik, 2003). Visual imagery may have an influential role herein, with the conjured images acting as relatively easy-retrievable semantic cues which could help the listener to better recollect one's music experience despite one's loss of self-awareness. Interestingly, rationality does not react strongly on setting evidence for low self-awareness (*low* for rationality stayed roughly the same; from 13% to 17%), suggesting that one's common sense is more or less impervious in case of impaired self-awareness. The network further suggests that during the listening experience, the thoughts that occurred were often reflective in nature (as opposed to having shallow thoughts), with its probability of *high* being 39%). Finally, further downstream the network, we observe that the high probability of experiencing visual imagery goes together with a milder increase in long-term memory (*high* from 37% to 48%), suggesting that imagery as expected is not always necessarily related to personal memories. In total, then, the network provides a first exploratory model which enables to examine the "flow" of information in a state of absorption and the function of imagery through a simplified and sparse structure.

Discussion

The network structure of phenomenal consciousness presented in Figure 16.1 shows that the main dimensions characterizing the experience of music absorption (i.e., full attention, altered experience, and loss of self-awareness) represent standard phenomenal attributes of consciousness. It provides useful insight into the structure of state absorption understood through the lens of a multidimensional perspective on consciousness. Likewise, it may also be helpful for future attempts to obtain an account of other music-induced states of subjective consciousness, including those that are described in Part III of this book. For instance, in order to understand what it would feel like to mind-wander in response to listening to music, the model could be *updated* (i.e., the probabilities are being re-estimated) after having provided appropriate evidence for those dimensions that are considered central to this form of inward distraction (Konishi, this volume). Likewise, the BN can also be used to further distinguish between different types of mind-wandering (e.g., *zoning out* and *tuning out*; Schooler et al., 2011) or absorption (*zoning in* and *tuning in*; Vroegh, 2019). Specifically, these zoning and tuning types are being differentiated in their degree of being concurrently aware of being in an actual state of mind-wandering or absorption. For these specific cases, then, we could additionally vary self-awareness to *low* or *high* levels to examine how their imposed intensity influence and spread out over the entire network differently. Finally, distinguishing between a negative-tinted ruminative state and a positive-tinted state of self-reflection (Morin, 2011) is also possible, yielding further insight into potential effects on the experience of visual imagery (Vroegh, 2019). Of course, an important but not unfeasible requirement would be to test the generalizability (external validity) of the network model in future research. For instance, the results for visual imagery in this study were, in terms of causes and effects, identical to the network found in the study where experimenter-selected music was used (Vroegh, 2021). The studies differed, though, in the probability for a high amount of imagery: for self-selected music 74%, for preselected music 45%.

What are the limitations that we have to keep in mind when interpreting the results of this analysis? It is important to emphasize that this network represents a single "snapshot" of consciousness averaged over time, used to model the phenomenal mind that is in an assumed equilibrium state of absorption. Hence, and inherent to using cross-sectional data, the presented model cannot demonstrate temporal relations between aspects of experience, although in reality experience develops and is maintained with both feedback and feedforward cyclical influences. Nevertheless, although these temporal effects were absent in this study, this does not diminish the usefulness and validity of this study in terms of an experiential model of music absorption.

Conclusions

When studying the structure of consciousness, focusing on its phenomenology is important, because "the features of the phenomenal level (how it is structured, how it dynamically changes across time, and so on) offer top-down constraints for the science of consciousness in the search for potential explanatory mechanisms in the brain" (Revonsuo, 2003, p. 3). In line with this notion and the recent trend towards explaining psychological phenomena by means of network approaches, the current chapter presented a probabilistic graphical model to study the network structure of music experience in response to a favourite, self-chosen piece of music. Specifically, it highlighted the conditional independences between visual imagery and other experiential features as part of being absorbed. An absorbed state of mind was found to be associated with a high probability of experiencing visual imagery (i.e., 74%); a finding that can be explained, at least in part, by using preferred, self-chosen music that is likely associated with various personal meanings and associated memories. Indeed, imagery was more profound compared to the study where experimenter-selected music was used (i.e., 45%, see Vroegh, 2021). In all, this study contributes to the small but growing amount of empirical research on music-induced imagery in the context of music absorption.

(Bayesian) network analysis is likely to hold promise in contributing to the field of the phenomenology of musical consciousness in the coming years. The recent availability of methods that allow for computing different types of networks and testing the replicability of parameter estimates will especially pay off when music researchers cross-compare subjective, first-person-perspective networks such as presented here with those obtained with "objective" brain-based data. Doing so will contribute to our understanding on the function and mechanisms of visual imagery in relation to music listening.

Notes

1 This chapter adheres to Taruffi and Küssner's (2019) definition of visual imagery as entailing a form of "seeing" images that are not activated by an external visual stimulus, but rather in relation to memory of past (usually personal) events related to specific times and places, abstract shapes like visual forms and/or colors, or else is a creative modification of previously obtained perceptual information.

2 Examples of items are, for *unity*: "The music and I remained clearly separated from each other.—I felt one with the music." *Reflective thoughts*: "I had superficial thoughts.—I was deeply concerned with profound questions or thoughts." *Mixed affect:* "I was filled with longing and nostalgia.—I felt no feelings of longing or nostalgia." *Long-term memory*: I was strongly reminded of something from my past.—I had no personal memories at all."

3 Note that altered experience is one of the main PCI dimensions, comprising the altered awareness of: (1) one's body image, (2) sense of time, (3) perception, and (4) transcendental meaning. It is akin to the dissociative feeling that emerges as part of an absorbing experience (see Carleton et al., 2010).

4 The data were resampled through bootstrapping, resulting in 5,000 networks of which subsequently the direction, significance, and strength of the edges were computed. This resulted in a single, averaged graph representing the frequency of edges that emerged in the bootstrapped repetitions and that is more robust against noisy data.

References

Agarwal, R., & Karahanna, E. (2000). Time flies when you're having fun: Cognitive absorption and beliefs about information technology usage. *MIS Quarterly*, 665–694. https://doi.org/10.2307/3250951

Becker, J. O. (2004). *Deep listeners: Music, emotion, and trancing*. Bloomington: Indiana University Press.

Borsboom, D., & Cramer, A. O. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, *9*, 91–121. https://doi.org/10.1146/annurev-clinpsy-050212-185608

Butler, L. D. (2004). The dissociations of everyday life. *Journal of Trauma and Dissociation*, *5*(2), 1–11. https://doi.org/10.1300/J229v05n02_01

Carleton, R., Abrams, M., & Asmundson, G. (2010). The Attentional Resource Allocation Scale (ARAS): Psychometric properties of a composite measure for dissociation and absorption. *Depression and anxiety*, *27*(8), 775–786. https://doi.org/10.1002/da.20656

Druzdzel, M. J. (1999). SMILE: Structural modeling, inference, and learning engine and GeNIe: A development environment for graphical decision-theoretic models (Intelligent Systems Demonstration). In *Proceedings of the Sixteenth National Conference on Artificial Intelligence and the Eleventh Innovative Applications of Artificial Intelligence Conference Innovative Applications of Artificial Intelligence* (pp. 902–903). Menlo Park, CA: American Association for Artificial Intelligence.

Farthing, G. W. (1992). *The psychology of consciousness*. Englewood Cliffs: Prentice-Hall.

Gabrielsson, A., & Wik, S. L. (2003). Strong experiences related to music: A descriptive system. *Musicae Scientiae*, *7*(2), 157–217. https://doi.org/10.1177%2F102986490300700201

Herbert, R. (2011). *Everyday music listening: Absorption, dissociation and trancing*. Farnham: Ashgate Publishing Limited.

Herbert, R. (2019). Absorption and openness to experience: An everyday tale of traits, states, and consciousness change with music. In R. Herbert, D. Clarke, & E. Clarke (Eds.), *Music and consciousness 2: Worlds, practices, modalities* (pp. 233–253). Oxford: Oxford University Press.

Huskey, R., Wilcox, S., & Weber, R. (2018). Network neuroscience reveals distinct neuromarkers of flow during media use. *Journal of Communication*, *68*(5), 872–895. https://doi.org/10.1093/joc/jqy043

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, *31*(5), 559–575. doi:10.1017/S0140525X08005293

Korb, K. B., & Nicholson, A. E. (2010). *Bayesian artificial intelligence* (2nd ed.). Boca Raton: Chapman & Hall

Küssner, M., & Eerola, T. (2019). The content and functions of vivid and soothing visual imagery during music listening: Findings from a survey study. *Psychomusicology: Music, Mind, and Brain*, *29*(2–3), 90–99. https://doi.org/10.1037/pmu0000238

Larsen, J. T., & McGraw, A. P. (2011). Further evidence for mixed emotions. *Journal of Personality and Social Psychology*, *100*(6), 1095–1110. https://doi.org/10.1037/a0021846

Morin, A. (2005). Possible links between self-awareness and inner speech theoretical background, underlying mechanisms, and empirical evidence. *Journal of Consciousness Studies*, *12*(4–5), 115–134.

Morin, A. (2011). Self-awareness part 1: Definition, measures, effects, functions, and antecedents. *Social and Personality Psychology Compass*, *5*(10), 807–823. https://doi.org/10.1111/j.1751-9004.2011.00387.x

Pekala, R. J. (1991). *Quantifying consciousness: An empirical approach*. New York: Plenum Press.

Pham, M. T. (2007). Emotion and rationality: A critical review and interpretation of empirical evidence. *Review of General Psychology*, *11*(2), 155–178. https://doi.org/10.1037%2F1089-2680.11.2.155

Rainville, P., & Price, D. D. (2003). Hypnosis phenomenology and the neurobiology of consciousness. *International Journal of Clinical and Experimental Hypnosis*, *51*(2), 105–129. doi:10.1076/iceh.51.2.105.14613

Revonsuo, A. (2003). The contents of phenomenal consciousness. *Psyche*, *9*(2), 1–12.

Schmittmann, V. D., Cramer, A. O. J., Waldorp, L. J., Epskamp, S., Kievit, R. A., & Borsboom, D. (2013). Deconstructing the construct: A network perspective on psychological phenomena. *New Ideas in Psychology*, *31*(1), 43–53. https://doi.org/10.1016/j.newideapsych.2011.02.007

Schooler, J. W., Smallwood, J., Christoff, K., Handy, T. C., Reichle, E. D., & Sayette, M. A. (2011). Meta-awareness, perceptual decoupling and the wandering mind. *Trends in Cognitive Sciences*, *15*(7), 319–326. https://doi.org/10.1016/j.tics.2011.05.006

Scutari, M., & Denis, J. B. (2015). *Bayesian networks: With examples in R*. Chapman and London: Hall/CRC.

Studerus, E., Gamma, A., & Vollenweider, F. X. (2010). Psychometric evaluation of the altered states of consciousness rating scale (OAV). *PLoS One*, *5*(8), e12412. https://doi.org/10.1371/journal.pone.0012412

Taruffi, L., & Küssner, M. B. (2019). A review of music-evoked visual mental imagery: Conceptual issues, relation to emotion, and functional outcome. *Psychomusicology: Music, Mind, and Brain*, *29*(2–3), 62–74. https://doi.org/10.1037/pmu0000226

Tellegen, A. (1981). Practicing the two disciplines for relaxation and enlightenment: Comment on role of the feedback in electromyograph feedback: The relevance of attention by qualls and sheehan. *Journal of Experimental Psychology*, *110*(2), 217–226. https://doi.org/10.1037/0096-3445.110.2.217

Tellegen, A. (1992). *Note on structure and meaning of the MPQ Absorption Scale* [Unpublished Manuscript]. Department of Psychology, University of Minnesota.

Tart, C. T. (1980). A systems approach to altered states of consciousness. In J. M. Davidson & R. J. Davidson (Eds.), *The psychobiology of consciousness* (pp. 243–270). New York: Plenum.

Vroegh, T. (2018). *The pleasures of getting into the music: Absorption and its role in the aesthetic appreciation of music* [Unpublished Doctoral Dissertation]. Goethe University, Frankfurt am Main.

Vroegh, T. (2019). Zoning in or tuning in? Identifying distinct absorption states in response to music. *Psychomusicology: Music, Mind, and Brain*, *29*(2–3), 156–170. https://doi.org/10.1037/pmu0000241

Vroegh, T. (2021). Visual imagery in the listener's mind: A network analysis of absorbed consciousness. *Psychology of Consciousness: Theory, Research, and Practice*. Advance online publication. https://psycnet.apa.org/doi/10.1037/cns0000274

Vuoskoski, J. K., & Eerola, T. (2015). Extramusical information contributes to emotions induced by music. *Psychology of Music*, *43*(2), 262–274. https://doi.org/10.1177%2F0305735613502373

17 Recumbent Journeys Into Sound—Music, Imagery, and Altering States of Consciousness

Jörg Fachner

Whatever Jimi Hendrix meant about being called from "my rainbow" in the lyrics of his 1967 recording *May this be love?* his phrase that "some people say that day-dreaming is for the lazy minded fools with nothing else to do" (Hendrix, 1967) points to a laidback, inwardly turned state of consciousness, in which letting the mind wander and allowing imagery and fantasy to take over is something that many may have enjoyed with mind-altering drugs in the summer of 1967. Guided Imagery and Music (GIM) has evolved out of a pharmaco-supported psychotherapy setting in which an altered state of consciousness (ASC)[1] was evoked with psychedelic drugs and participants were listening to extended playlists of classical or jazz music (Bonny & Pahnke, 1972; Dukić, this volume). However, Helen Bonny, the founder of GIM, observed that inducing an ASC with deep relaxation methods and a subsequent guided journey of the "travellers" (patients) in a reclined or supine position elicited sufficient imagery for the therapy session without the psychedelics (Bonny, 1980). This may raise the question of which role the combination of recumbent body postures and ASC may have for the imagery that occurs along with the music listened to. In this chapter, we aim to focus on how the occurrence of mental imagery during music listening in different ASC-related "imagery settings" is enhanced by assuming a recreational recumbent posture. Although mental imagery has different sensory domains (visual, motor, kinaesthetic, tactile, and olfactory; see Part I of this volume), the imagery settings here will exemplify the visual imagery experienced in more detail (see Taruffi & Küssner, this volume). The main argument is that a recumbent body posture and less movement during an "ecstatic" ASC as Rouget (1985) may call it induces a trophotropic downregulation in order to free up energy for focusing on the imagery occurring along with the music attended to.

Altering Sound and Vision

Deviations from the normal states of consciousness have been classified by Fischer (1998) and classified on a psychophysiological continuum between ergotropic (excitatory/"expenditure of energy"[2]) and trophotropic (relaxing/"renewal of energy") states. According to Rouget (1985), trance (lat., transitio) and ecstasy (lat., ecstasis, out of one's head) are the most common deviations from

DOI: 10.4324/9780429330070-21

normal waking consciousness in connection with music. *Trance* seems to have a more ergotropic profile with high vigilance states and extended movement, while *ecstasy* seems to be related to trophotropic downregulation, lower vigilance states, less movement, and increased mental activity and imagery (Fachner, 2011). An inquiry on out-of-body experiences showed that they occur more often in immobility, when lying down supine or sitting relaxed (Zingrone et al., 2010), that is, when the focus of attention can turn inward, and more afferent information is processed. Dittrich (1998), in his international comparison of altered states of consciousness finds that, independent of the stimulus, there are five core ASC experiences: “Oceanic Boundlessness (OSE),” “Dread of Ego Dissolution (AIA),” “Visionary Restructuralization (VUS),” “Auditive Alteration (AVE),” and “Vigilance Reduction (VIR).”

Set and Setting: Recumbent Journeys into Ecstasy, Sound, and Imagery

The following settings investigating the “traveller’s” inner “ecstatic” journey into sound and imagery are characterized by a recumbent immobility, that is, a state of contemplation and inwardly turned attention. The examples were chosen to exemplify how increased imagery and recumbent journeys into sound are related in different ASC induction settings. The common denominator is that in all five settings, music is heard in a recumbent position and that these settings are used to enhance the imagery during listening. The *set*, as the individual psychophysical condition in connection with personal memories, attitudes, and experienced temporality of events, and the *setting*, as the actual physical, social, and symbolic environment in which events happen, play an important role in triggering ASC and contextualizing body posture and movement (Carhart-Harris et al., 2018; Fachner, 2011). A recumbent position and closed eyes in order to drift away into the realm of imagery while lying on a couch and being in a room with dimmed light is what also did signify the social space in which opium was consumed (Jonnes, 1999). Those “travelling” into their inner worlds of ecstasy are paradoxically not moving much in a laid-back or recumbent eyes closed condition. Sigmund Freud already noticed the importance of a recumbent body posture for his free association part of psychoanalysis. It was easier for the patient to freely associate when lying relaxed with eyes-closed on a couch or sitting recumbent in a big armchair.

Imagery Setting I: Shamanic Journeys

Flor-Henry et al. (2017, p. 5) define shamanic states of consciousness as follows:

> Shamanic States of Consciousness are a distinct sub-category of ASCs characterized by lucid but narrowed awareness of physical surroundings, expanded inner imagery, modified somatosensory processing, altered sense of self, and an experience of spiritual travel to obtain information necessary for solving a particular individual or social problem.

The focus here on "expanded inner imagery" is substantiated by a rich case study of a shamanic journey recorded with an eyes-closed EEG recording in immobility. The shaman describes her ASC experience as follows:

> "An inner vision of a wolf face, very close to my face, came soon after. Then an owl (two close representations are joined). I have to add that these visions were not seen as precisely as they are with our eyes. It may happen that the visions are quite similar to how we see with our eyes, but sometimes it's more like the feeling of a presence than the image of it." . . . "My eyes (closed during every trance) have visions like geometrical patterns, places or entities called spirits by the shamans. They can appear to me as men, or women, or animal patterns, that I can suddenly speak with." . . . "I suddenly feel like the animal or the entity I am supposed to be possessed by. As I am a wolf, for example, I can feel paws in place of my hands and a muzzle in place of my nose. I don't speak anymore but howl like a wolf."
>
> (Flor-Henry et al., 2017, pp. 16–17)

Flor-Henry's interpretation of the corresponding EEG is based on lateralization patterns and topographic state changes compared to a normative database. In the shamanic state of consciousness, a right and rear shift resulting in a predominantly right posterior sensory mode of inwardly turned attention. This resembles Dietrich's (2003) hypothesis of a hypo-frontal state during ASC, in which attention turns inward, and is dominating the brain activity.

Imagery Setting II: Monotonous Drumming

Kjellgren and Eriksson (2010) studied 22 people listening to live monotonous drum music in a darkened room for 20 minutes. Lively visual imagination and physical sensations associated with the drumbeat reinforced the impression of actually travelling and encounters with landscapes on the journey including people, animals, and plants.

In another study, 15 experienced "travellers" heard an eight-minute monotonous drum piece in the fMRI (Hove et al., 2016). In the ASC condition, Michael Harner's CD *Solo Drumming* was played and a corresponding shamanic journey into the underworld was undertaken, while in the control condition, the requirement was not to go on such a journey and to listen only. In addition, during the control condition, the drum pulse was a bit more irregular; the deviation from the average beat to beat inter-onset interval was 16.5 ms instead of 6.5 ms in ASC. During ASC, but not during the control condition, three distinct brain regions, relevant for maintaining inwardly turned cognitive attention, were increasingly connected with each other, resulting in a stronger focus on internal processes.

Apparently, the monotonous repetitive nature of drumming is what inhibits the auditory activation, by decoupling the auditory input from inner activity that increases attention to inner imaginative processes. The measurements of functional connectivity of particular brain regions linking with the auditory cortex were significantly reduced during the trance. It is precisely the predictability of regular drumming at a stable pace that facilitates the “work” and shutdown of the brain, which can be demonstrated by a corresponding decrease in the N1 amplitude in the brain wave spectrum of the EEG after repeated isochronous stimulation (Will & Makeig, 2011, p. 104).

It seems that listening to monotonous drumming (or any other monotonous droning sound) downregulates brain activity, that is, diminishes energy consumption of auditory processing and uses the released capacity for deepening imagery and journeying. In this study, the (monotonous) music acts as a means for switching off the outside and focusing on the inner world. Neuronal networks in the brain temporarily reconfigure (Hove et al., 2016) and one goal of the reconfiguration strategy during ASC is a decoupling of external sensory impressions from the cognitive processing of internal processes (perceptual decoupling). This state makes it possible to shut-out the outside world and focus primarily on the increasing inner images, streams of thoughts and feelings, even when lying in an MRI scanner or sitting relaxed during EEG-recordings.

Imagery Setting III: Techno-Shamanic Diffusion

Although the everyday connotation of the terms *trance* and *ecstasy* may have diametrical or similar meanings when connected to music, in the so-called “Techno” music genre, *trance* still stands for dance and excitation and *ecstasy* refers to a meditative “chill-out” music, representing the “laidback” relaxation state after exhaustive dancing (Hutson, 2000; Penman & Becker, 2009; St. John, 2011).

Following this distinction of trance and ecstasy, which links the effect of music to the extent of physical activity and subdivides it into (physically) active and inactive states of consciousness, Rouget’s (1985) understanding of ecstasy as a journey into the inner worlds was expanded into the concept of “shamanic diffusions” with regard to a setting of electroacoustic music concerts within which the (chill-out) recumbently listing audience is immersed into sound. Here, “the musical experience then facilitates a journey through illusory sonic environments” which then is also related to the “solitary, immobility of the electroacoustic listening experience . . . as shamanic journeys or rituals through the diffusion of sound” (Weinel, 2014, p. 5).

To realize this, Weinel has produced a series of compositions and electroacoustic sound treatment procedures (Weinel, 2011, 2012) that would allow a similar experience of sound alterations as for example under the influence of psychedelic drugs. These sound alterations of electro-acoustic music are based on users descriptions of ASC and auditory changes as held in the EROWID library (Weinel et al., 2014) or to some extent already systematized in the book *On Being Stoned* (Tart, 1971).

However, to illustrate how such shamanic diffusions maybe realized, Weinel (2014, p. 6) writes:

> For example, during a hallucinatory experience one may perceive distortions to time-perception or spatial awareness. One may also see visual patterns of

> hallucination, hear auditory illusions or imagine strange entities. These can be used as a basis for the design of sonic materials and spaces. For example: if a hallucination contains whispering voices in the context of a room that seems to grow and shrink unnaturally, corresponding sonic materials can be produced and spatial processes can be applied. Likewise, hallucinatory experiences often have a gradual onset, plateau and termination. These can be used as a basis for the design of corresponding musical sections.

Imagery Setting IV: Ritualistic Use of Ayahuasca and Iboga

Katz and De Rios (1971) explained the function of the "Icarus" songs, sung and whistled in the Peruvian Ayahuasca ceremonies, as being to help the shaman and also their clients to control the drug action. They compared music's function to a "jungle gym," giving a structure to control drug-induced ASC, and providing "a series of paths and banisters to help them negotiate their way" (De Rios & Janiger, 2003, p. 161). The music and its inner structure serve to provide a functional replacement and support for the psychic structure in drug-related phases of ego dissolution (Carhart-Harris & Friston, 2010), not only to create a certain mood within the ritual setting.

The co-stimulation of different senses (so-called *synaesthesia*) seems to be an important part of rituals with drugs. Hallucinogens such as Ayahuasca or LSD can cause sensory interactions of sensory modalities such as smelling, seeing, feeling, touching, hearing, and tasting, thus leading to increased consistency of perceptual content and form or bringing visions to life (Baudelaire, 1860; Shanon, 2011; Sinke et al., 2012).

> When music is present, it usually guides the visions. In particular, the tempo and rhythm of the music people hear is often reflected in their visions. Notably, the music determines the pace and movement of figures appearing in the visions as well as the rate of change between images. Some note that their visions last as long as the singing, and when the singing stops, so do the visions.
> (Shanon, 2011, p. 286)

The interaction of music and drug can be used for ritual-immanent purposes, that is, to use change of tempo or certain rhythms to mark particular sections during the rituals (Fachner, 2011; Katz & De Rios, 1971; Maas & Strubelt, 2006; Shanon, 2011).

Maas and Strubelt (2006) studied the music and Iboga-pharmacotherapy of traditional healers in Gabon, Africa. During the three-day ritual, polyrhythmic music was heard continuously, and according to Maas' Iboga experience, a persistent "inner metre" or rhythmic pulsation set in. Among several observations exemplified, under the Iboga's influence, the geometric patterns in the surrounding artworks and paintings seem to visualize the rhythmic patterns perceived in the music and the ritual meaning of polyrhythms was understood analogously to the interweaving of different paths and intersections of life paths.

Imagery Setting V: LSD, Peak Experiences, and Psychotherapy/ Psychiatric Research

In psychiatry, model-psychosis research served to compare psychotic states of hallucinations with drug-induced hallucinations and to discuss its noetic and clinical considerations (Gouzoulis-Mayfrank et al., 1998; Leuner, 1962). Aims were to describe hallucinogenic states like the productive states of schizophrenia, which seem to be analogous to some experiences made during psychedelic drug action.

The psychiatrist Osmond introduced the term *psychedelic*, which means "expanding psychic reality" (Tanne, 2004). Founders of psychedelic therapy discovered that a setting, realized via works of art, fairy tales, narratives, music, ritual and selected interior, had a profound capacity to stimulate the unconscious and to provoke associations to yield unconscious material for the therapy (Barrett et al., 2018; Fachner, 2007; Grof, 1994; Leuner, 1962; Melechi, 1997). The psychedelic agent intensified the experience, changed the perception levels of stimulation modes, and weakened defence mechanisms.

Psychedelics evoke "peak experiences in the right people under the right circumstances" as Abraham Maslow (1970, p. 27) described in his book on values and *peak experiences* pointing to an important concept of psychedelic therapy being interested to force peak experiences for gaining insight and allow a vast amount of psychic material for the psychotherapy setting by optimizing the environment for the drug experience.

Guided imagery and Music (GIM) has evolved out of "psychedelic peak therapy" (Bonny, 1980; Bonny & Pahnke, 1972; Dukić, this volume) in which certain pieces of mostly classical or jazz music were played in a thematic therapeutic sequence to facilitate emotions, evoke peak experiences, uncensored responses, and associations, and to open a path to the inner world of the client's unconscious. All these happened in a relaxed secure and guided setting of psychedelic therapy. Anti-toxicants for a possible bad trip were at hand and therefore the patient could let go (Bonny & Pahnke, 1972; Carhart-Harris et al., 2018).

However, Helen Bonny observed that shorter therapy sessions (the LSD sessions lasted 8–10 hours) inducing an ASC with deep relaxation methods (without LSD) and a subsequent journey in a reclined or supine position elicited sufficient imagery for the therapy session (Bonny, 2002, pp. 144–145). Nevertheless, recent research into Psilocybin, LSD, and music listening (Kaelen et al., 2015; Kaelen et al., 2018) showed that the vividness of imagery is stronger under the influence,[3] and one reason is the breakdown of ego defence mechanisms (Carhart-Harris & Friston, 2010; Kaelen et al., 2016). The emergence of important imagery along with the music yield strong imagery patterns with a noetic structure that may well bring up important psychic material that can be used for therapy (Bonny, 1980). Imagery enhancement under the influence is signified by increasing functional connectivity between para-hippocampus (PHC) and visual cortex (VC), and furthermore, PHC to VC information flow in the interaction between music and LSD was observed. The latter result "correlated positively with ratings of enhanced eyes-closed visual imagery, including imagery of an autobiographical nature" (Kaelen et al., 2016, p. 1100).

The ASC experience yielded with the relaxation induction methods can be controlled voluntarily and—importantly—in an interaction between a therapist and a patient (Fachner et al., 2019) while the imagery intensity under the influence may put the traveller in a "floodlight state" that needs longer time to be processed and integrated (Fachner, 2007; Grocke, 1999; Melechi, 1997). Research into deep relaxation with music and imagery (Pfeifer et al., 2016) and reviewing common features of music and altered states point out that setting, performance rites, suggestibility traits, and personal willingness to go into an altered state are of importance for the meaning of the imagery (Fachner, 2011). In a study on emotional peak experiences (without drugs), the participants reported stronger emotional experiences or mystical images in music at a faster pace than in quiet and slow pieces (Lowis, 1998), indicating music-inherent triggers of peak experiences.

Conclusion

We have seen how a recumbent setting that promotes an increase in inner imagery is related to the evocation of differing forms of ASC. In ecstatic healing settings, instances of immobility and the intended altering of the focus of attention to the inner world are used to facilitate different imaginations of one-self, in order to vary perspective and to allow introspection. No matter how the induction of the journey is seated, the important part for the healing setting is that meaningful and hidden material can be brought to the forefront of cognitive and emotional processing. A guided imagery setting with an experienced healer who accompanies the journey and knows the music can promote change of the time course of the imagery and narratives related to personal meaning, and can enhance emotional intensity and relevance of the imagery (Dukic et al., 2021; Fachner et al., 2019).

Music with differing degrees of complexity and especially monotonous structures may help to un-focus from the outer world and to be brought into a "jungle gym" of cross-modal correspondences that may vary individually, but seem to be guided by tempo, mode, and intensity of musical structure in order to evoke an imaginative equivalent of visual form constants and structural movements. This process is culturally mediated and may serve different noetic horizons of meaning which may be evoked with association patterns learned or mediated via symbolic structures co-occurring in the setting of the imagery.

However, in the ASC state of ecstasy, it is of vital importance that a setting is realized in which the brain can free energy for the cross-modal imagery processes. Psychedelic drugs seem to enhance cross-modal imagery processes and meaningful identification of layers and correspondences. Whether the music is monotonous or complex to induce rich imagery depends on the setting, but the reduced amount of body movement seems to be important for free energy. Further research into settings that promote these journeys is needed to describe its general constituents.

Notes

1 Ludwig (1966) describes the following "general characteristics of altered states of consciousness": alterations in thinking, disturbed time sense, loss of control, change in emotional expression and body image, perceptual distortions, change in meaning or significance, sense of ineffable, feelings of rejuvenation, and hyper suggestibility.
2 Compare APA definitions of ergo- and trophotropic here: American Psychological Association (n.d.).
3 Helen Bonny in a personal meeting in November 1999 with the author of this chapter also confirmed that the psychedelics brought up more and stronger imagery in the psychedelic therapy sessions (H. Bonny, personal communication, November 1999).

References

American Psychological Association. (n.d.). *APA dictionary of psychology*. Retrieved October 6, 2021, from https://dictionary.apa.org/

Barrett, F. S., Preller, K. H., & Kaelen, M. (2018). Psychedelics and music: Neuroscience and therapeutic implications. *International Review of Psychiatry*, *30*(4), 350–362. https://doi.org/10.1080/09540261.2018.1484342

Baudelaire, C. (1860). *Les paradis artificiels: Opium et haschisch*. Paris: Poulet-Malassis et de Broise.

Bonny, H. (1980). *GIM therapy: Past, present and future implications* (Vol. 3). Salina: The Bonny Foundation.

Bonny, H. (2002). *Music and consciousness: The evolution of guided imagery and music*. Gilsum: Barcelona Publishers.

Bonny, H. L., & Pahnke, W. N. (1972). The use of music in psychedelic (LSD) psychotherapy. *Journal of Music Therapy*, *9*(2), 64–87. https://doi.org/10.1093/jmt/9.2.64

Carhart-Harris, R. L., & Friston, K. J. (2010). The default-mode, ego-functions and free-energy: A neurobiological account of Freudian ideas. *Brain*, *133*(4), 1265–1283. https://doi.org/10.1093/brain/awq010

Carhart-Harris, R. L., Roseman, L., Haijen, E., Erritzoe, D., Watts, R., Branchi, I., & Kaelen, M. (2018). Psychedelics and the essential importance of context. *Journal of Psychopharmacology*, *32*(7), 725–731. https://doi.org/10.1177/0269881118754710

De Rios, M. D., & Janiger, O. (2003). *LSD, spirituality, and the creative process*. Rochester: Park Street Press.

Dietrich, A. (2003). Functional neuroanatomy of altered states of consciousness: The transient hypofrontality hypothesis. *Consciousness and Cognition*, *12*(2), 231–256. https://doi.org/10.1016/S1053-8100(02)00046-6

Dittrich, A. (1998). The standardized psychometric assessment of Altered States of Consciousness (ASCs) in humans. *Pharmacopsychiatry*, *31*(2), 80–84. https://doi.org/10.1055/s-2007-979351

Dukic, H., Parncutt, R., & Bunt, L. (2021). Narrative archetypes in the imagery of clients in guided imagery and music therapy sessions. *Psychology of Music*, *49*(2), 287–303. https://doi.org/10.1177/0305735619854122

Fachner, J. (2007). "Takin' it to the streets . . ." Psychotherapie, Drogen und Psychedelic Rock: Ein Forschungsüberblick. *Samples, Online-Publikationen des Arbeitskreis Studium Populärer Musik e.V.*, *6*. http://aspm.ni.lo-net2.de/samples/6Inhalt.html

Fachner, J. (2011). Time is the key: Music and ASC. In E. Cardeña & M. Winkelmann (Eds.), *Altering consciousness: A multidisciplinary perspective: Vol. 1. History, culture and the humanities* (pp. 355–376). Santa Barbara: Praeger.

Fachner, J., Maidhof, C., Grocke, D., Nygaard Pedersen, I., Trondalen, G., Tucek, G., & Bonde, L. O. (2019). "Telling me not to worry . . ." hyperscanning and neural dynamics of emotion processing during guided imagery and music. *Front Psychol*, *10*(1561). https://doi.org/10.3389/fpsyg.2019.01561

Fischer, R. (1998). Über die Vielfalt von Wissen und Sein im Bewusstsein. Eine Kartographie außergewöhnlicher Bewusstseinszustände. In R. Verres, H. C. Leuner, & A. Dittrich (Eds.), *Welten des Bewußtseins* (pp. 43–70). Berlin: VWB-Verlag für Wissenschaft und Bildung.

Flor-Henry, P., Shapiro, Y., & Sombrun, C. (2017). Brain changes during a shamanic trance: Altered modes of consciousness, hemispheric laterality, and systemic psychobiology. *Cogent Psychology*, *4*(1), 1313522. https://doi.org/10.1080/23311908.2017.1313522

Gouzoulis-Mayfrank, E., Hermle, L., Thelen, B., & Sass, H. (1998). History, rationale and potential of human experimental hallucinogenic drug research in psychiatry. *Pharmacopsychiatry*, *31* (S2), 63–68. https://doi.org/10.1055/s-2007-979348

Grocke, D. (1999). *A phenomenological study of pivotal moments in Guided Imagery and Music (GIM) therapy* [Doctoral Dissertation, University of Melbourne]. University of Melbourne, University Library. http://hdl.handle.net/11343/39018

Grof, S. (1994). *LSD psychotherapy* (2nd ed.). Alameda: Hunter House.

Hendrix, J. (1967). *Are you experienced?* [Long Player]. London: Polydor International.

Hove, M. J., Stelzer, J., Nierhaus, T., Thiel, S. D., Gundlach, C., Margulies, D. S., . . . Merker, B. (2016). Brain network reconfiguration and perceptual decoupling during an absorptive state of consciousness. *Cerebral Cortex*, *26*(7), 116–3124. https://doi.org/10.1093/cercor/bhv137

Hutson, S. R. (2000). The rave: Spiritual healing in modern western subcultures. *Anthropological Quarterly*, *73*, 35–49. www.jstor.org/stable/3317473

Jonnes, J. (1999). *Hep-cats, narcs and pipe dreams* (1st ed.). Baltimore: Johns Hopkins University Press.

Kaelen, M., Barrett, F. S., Roseman, L., Lorenz, R., Family, N., Bolstridge, M., . . . Carhart-Harris, R. L. (2015). LSD enhances the emotional response to music. *Psychopharmacology*, *232*(19), 3607–3614. https://doi.org/10.1007/s00213-015-4014-y

Kaelen, M., Giribaldi, B., Raine, J., Evans, L., Timmerman, C., Rodriguez, N., . . . Carhart-Harris, R. (2018). The hidden therapist: Evidence for a central role of music in psychedelic therapy. *Psychopharmacology*, *235*(2), 505–519. https://doi.org/10.1007/s00213-017-4820-5

Kaelen, M., Roseman, L., Kahan, J., Santos-Ribeiro, A., Orban, C., Lorenz, R., . . . Carhart-Harris, R. (2016). LSD modulates music-induced imagery via changes in parahippocampal connectivity. *European Neuropsychopharmacology*, *26*(7), 1099–1109. https://doi.org/10.1016/j.euroneuro.2016.03.018

Katz, R., & De Rios, M. D. (1971). Whisteling in peruvian ayahuasca healing sessions. *Journal of American Folklore*, *84*(333), 320–327. https://doi.org/10.2307/539808

Kjellgren, A., & Eriksson, A. (2010). Altered states during shamanic drumming: A phenomenological study. *International Journal of Transpersonal Studies*, *29*(2), 1–10. https://doi.org/10.24972/ijts.2010.29.2.1

Leuner, H. C. (1962). *Die experimentelle Psychose* [The experimental psychosis]. Berlin: Springer.

Lowis, M. J. (1998). Music and peak experiences: An empirical study. *Mankind Quart*, *39*(2), 203–224.

Ludwig, A. M. (1966). Altered states of consciousness. *Archives of General Psychiatry*, *15*(3), 225–234. doi:10.1001/archpsyc.1966.01730150001001

Maas, U., & Strubelt, S. (2006). Polyrhythms supporting a pharmacotherapy: Music in the Iboga initiation ceremony in Gabon. In D. Aldridge & J. Fachner (Eds.), *Music and altered states: Consciousness, transcendence, therapy and addictions* (1st ed., pp. 101–124). London: Jessica Kingsley.

Maslow, A. (1970). *Religions, values and peak experience*. New York: Viking Penguin.

Melechi, A. (Ed.). (1997). *Psychedelia Britannica* (1st ed.). London: Turnaround.

Penman, J., & Becker, J. (2009). Religious ecstatics, "deep listeners," and musical emotion. *Empirical Musicology Review*, *4*(2), 49–70.

Pfeifer, E., Sarikaya, A., & Wittmann, M. (2016). Changes in states of consciousness during a period of silence after a session of Depth Relaxation Music Therapy (DRMT). *Music & Medicine*, *8*(4), 180–186. https://doi.org/10.47513/mmd.v8i4.473

Rouget, G. (1985). *Music and trance: A theory of the relations between music and possession*. Chicago: University of Chicago Press.

Shanon, B. (2011). Ayahuasca and music. In E. Clarke & D. Clarke (Eds.), *Music and consciousness* (pp. 281–294). Oxford: Oxford University Press.

Sinke, C., Halpern, J. H., Zedler, M., Neufeld, J., Emrich, H. M., & Passie, T. (2012). Genuine and drug-induced synesthesia: A comparison. *Consciousness and Cognition*, *21*(3), 1419–1434. https://doi.org/10.1016/j.concog.2012.03.009

St. John, G. (2011). Spiritual technologies and altering consciousness in contemporary counter culture. In E. Cardena & M. Winkelman (Eds.), *Altering consciousness: Multidisciplinary perspectives: Vol. 1. History, culture, and the humanities* (1st ed., pp. 203–228). Santa Barbara: Praeger.

Tanne, J. H. (2004). Humphry Osmond. *BMJ: British Medical Journal*, *328*(7441), 713–713. www.ncbi.nlm.nih.gov/pmc/articles/PMC381240/

Tart, C. (1971). *On being stoned: A psychological study of marihuana intoxikation* (1st ed.). Palo Alto: Science and Behaviour Books.

Weinel, J. (2011, July 31–August 5). *Tiny jungle: Psychedelic techniques in audio-visual composition* [Paper presentation]. 2011 International Computer Music Conference (ICMC), Huddersfield, UK.

Weinel, J. (2012). *Altered states of consciousness as an adaptive principle for composing electroacoustic music* [Doctoral Dissertation, Keele University London South Bank University]. Aalborg Universitet. https://vbn.aau.dk/en/publications/altered-states-of-consciousness-as-an-adaptive-principle-for-comp

Weinel, J. (2014). Shamanic diffusions: A technoshamanic philosophy of electroacoustic music. *Sonic Ideas/Ideas Sonicas*, *6*(12).

Weinel, J., Cunningham, S., & Griffiths, D. (2014, October 1–3). *Sound through the rabbit hole: Sound design based on reports of auditory hallucination* [Paper presentation]. 9th Audio Mostly: A Conference on Interaction With Sound, Aalborg, Denmark.

Will, U., & Makeig, S. (2011). EEG research methodology and brainwave entrainment. In J. Berger & G. Turow (Eds.), *Music, science, and the rhythmic brain* (1st ed., pp. 85–110). London: Routledge.

Zingrone, N. L., Alvarado, C. S., & Cardena, E. (2010). Out-of-body experiences and physical body activity and posture: Responses from a survey conducted in Scotland. *The Journal of Nervous and Mental Disease*, *198*(2), 163–165. doi:10.1097/NMD.0b013e3181cc0d6d

Part IV

Applied Mental Imagery

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

18 Imagery and Movement in Music-Based Rehabilitation and Music Pedagogy

Rebecca S. Schaefer

Mental imagery, or the human ability to experience perception or actions without sensory stimulation or overt movements, is highly common, and generally considered to be part of our daily lives.[1] Mental imagery arguably supports several other cognitive functions such as memory processes (Baddeley & Andrade, 2000; Navarro Cebrian & Janata, 2010; Schaefer, 2018), emotional processes (Wicken et al., 2019), and various kinds of learning (Halpern & Overy, 2019; Paivio, 1969; Tartaglia et al., 2009), even in non-human animals (Blaisdell, 2019). As such, imagery might be harnessed as a valuable resource for practical settings in which simulation of sound, images, or movement (or some combination of these) is helpful, and its use may be improved by increasing understanding of the underlying mechanisms.

Mental imagery is considered here as a representation that may only involve one sense (or modality), termed unimodal imagery, or include multiple senses (such as a visual image combined with sound, or touch, or movement), termed multimodal imagery. Additionally, distinctions can be made in the level of intentionality, referring to the extent to which imagery is effortfully and intentionally produced as a deliberate cognitive action, or spontaneously emerges. The distinctions in modality or level of intentionality likely do not have hard boundaries, but rather represent spectra, meaning that an image may for instance be predominantly auditory but include some vague sense of movement, or may be initiated intentionally at first but then continue without any effort. For the current discussion, the focus is on effortful, conscious imagery applied as a tool in learning processes.

The current chapter focuses specifically on how imagery can support movement and movement learning. To this end, two main application areas will be considered: movement rehabilitation and music pedagogy. While the functional goals in these two domains are very different (i.e., to recover "normal" movement function vs. to attain highly skilled, specialized movement), both processes arguably rely on learning mechanisms that support motor skill acquisition (Maier et al., 2019; Sampaio-Baptista et al., 2018). Both thus likely involve forming new or renewed movement representations in the motor system of the brain, implemented by developing novel learning-related neural connections. In both cases, repetition and the salience of the movement experience, among others, are good predictors of movement acquisition (e.g., Kleim & Jones, 2008; Korman et al.,

DOI: 10.4324/9780429330070-23

2003). Although studies are sparse, there are indications that methods using mental imagery represent a low-cost, adaptable way to support movement learning as well as music pedagogy, albeit using different elements of imagery (i.e., through its time structure when cueing movement, vs. the association with an emotional quality for expressivity, or multimodal integration for reduced cognitive load).

In the following, the use of mental imagery for motor learning is considered, focusing first on how music imagery might replace or support the use of auditory cues in rehabilitation, followed by how other aspects of imagery are used in music pedagogy.

Music Imagery in Motor Rehabilitation

Musical rhythms are increasingly used as auditory cues in movement rehabilitation for a range of disorders that include motor impairments, such as Parkinson's disease (PD), stroke, Huntington's disease (HD), and various others (Schaefer, 2014a; Thaut, 2005; Thaut et al., 2015), focusing specifically on rehabilitation exercises based on recurring, periodic movements, aligned to music. There are several potential explanations on how auditory rhythms may support movement and movement re-learning processes. Based on work with PD patients specifically, moving to auditory rhythm is thought to recruit a neural pathway differing from moving without sound (namely utilizing more of premotor and parietal areas rather than presupplementary and supplementary motor areas; Nombela et al., 2013, see also Schaefer & Overy, 2015). The underlying mechanisms of how auditory cues might alter movement learning have been described as four potential, non-mutually exclusive phenomena (Schaefer, 2014a), namely more stable periodic movement (facilitated through repetitive patterns in music (Margulis, 2013); qualitatively different movement implementation due to the perceptual embedding of the auditory cue, as supported by findings reported by Moore et al. (2017); supporting temporal processing more generally (as suggested by the carry-over effects of cued music to for instance the perceptual domain; Benoit et al., 2014); and increased motivation and enjoyment during practice, such as for instance reported in Moumdjian et al. (2019). This, however, begs the question whether imagined music might reliably take the place of heard music.

Interestingly, some studies anecdotally mention that rehabilitation programmes using music also often lead to spontaneous imagery of the music that was used, and spontaneous use of imagery to sustain regular movement (e.g., Schauer & Mauritz, 2003), in line with emerging knowledge on the relationship between spontaneous music imagery and movement, particularly that regular movement appears to more often induce spontaneous music imagery (Campbell & Margulis, 2015; Floridou et al., 2015). Additionally, positive results based on other forms of self-cueing are reported, such as singing (Harrison et al., 2017), or even inner singing (Harrison et al., 2019; Satoh & Kuzuhara, 2008) to regularize one's own gait. This suggests that using music imagery to self-cue movement is not only a viable task, but it may also be effective in rehabilitation settings.

The neural correlates of music imagery as a cue for movement were investigated for wrist flexion movements, comparing moving to imagined music to moving without imagery, as well as to moving to heard music (Schaefer et al., 2014). Here, activation in the supplementary motor area (SMA), more commonly reported for music imagery (Halpern & Zatorre, 1999; Herholz et al., 2012), was also seen while participants moved, in addition to activation related to that movement itself. This suggests that this SMA activation, when seen in the absence of movement, is not related to covert movement (Halpern & Zatorre, 1999) but rather, as also suggested elsewhere (cf. Halpern et al., 2004) related to the actual task of imagining music, likely including functions such as tracking and sequencing events over time internally. Additionally, activation was seen in the basal ganglia for imagery-cued movement that may be related to initiating movement (Schaefer et al., 2014); however, this needs further investigation. In the electro-encephalogram (EEG), moving the fingers to imagined or actually heard music showed a similar effect when compared to moving in silence (Floridou et al., 2018). Here, the finding of similar neural responses for both heard and imagined musical cues as compared to silence further supports the notion of additional activity in motor networks when movement is rhythmically cued. How this additional activation may support learning processes remains an open question, but these initial studies suggest that while imagined music may not affect movement identically to heard music, movement timing measurements indicate that music imagery can function as a self-generated cue for movement (cf. Floridou et al., 2018). However, more research is needed to assess the impact of self-generated cues on movement learning over a longer period of time. Moreover, further inquiry into individual differences into imagery abilities, specifically the vividness and potentially stability of the imagined pulse, and how these abilities may be affected in different clinical populations, is crucial to determine the usefulness of imagined auditory cues in clinical contexts.

Multimodal Imagery in Music Pedagogy

Imagery has been described as a crucial element of musical processes generally, with auditory imagery abilities (such as vividness) explicitly featuring in early tests of musical aptitude (i.e., Gordon, 1965; Seashore, 1915), and understanding and appreciation of music being argued to also be based on broad inner simulation and imagery skills, albeit more implicitly (Hargreaves, 2012). In the context of music pedagogy, some music training methods have expressly focused on developing "inner hearing," where representations of musical structure are trained and rehearsed in the absence of heard sound (see examples of Kodály's methods described in Halpern & Overy, 2019), or even more explicitly when imagined situations or scenes are used as exercises for musicians, intended to increase imagery skills, and possibly even musical creativity (Adolphe, 2013). Additionally, for performance and motor control aspects such as timing, subconscious imagery processes are argued to be fundamental to solo or joint music-making (Keller, 2008, 2012), increasing the potential to interact with a musical partner by mentally simulating their sounds and perhaps movements, and better anticipate what will come next. Importantly, in these different scenarios, different kinds of imagery are at play. In terms of modality, imagery here likely involves auditory, movement, and possibly other sensory domains such as visual or haptic characteristics (Godøy, this volume). In terms of intentionality and conscious experience, it is something different to simulate sound or movement effortfully, as in mental practice, or using predictive mechanisms to increase anticipatory behaviour, which—although arguably based on mental representations of actual sound or movement—tends to happen subconsciously (cf. Schaefer, 2014b). When considering applications of imagery techniques in music pedagogy, again intentional imagery, which can be deliberately used as a learning strategy, is arguably most easily adapted into a teaching method. Halpern and Overy (2019) put forward the idea that specifically the additional activation found in SMA, as compared to actual perception, may be an indication that imagining music may involve a deeper level of processing or concentration than listening alone, and thus wields a stronger potential for learning.

One example of imagery processes being used to facilitate better piano performance was provided by Davidson-Kelly et al. (2015), who describe a method intended to facilitate memorization of large piano concertos. Designed by Nelly Ben-Or, a concert pianist and Alexander Technique expert, this method focuses on practice away from the instrument, internalizing auditory, motor, and visual aspects (i.e., the score), creating what she terms a "Total Inner Memory." This multimodal representation of the score, when securely created, arguably frees up cognitive resources during performance by reducing heavy memory load. This allows more focus on expressiveness in motor control, as well as quick access to the piece in the future (more detailed descriptions provided in Davidson-Kelly et al., 2015).

Another example of how deliberate imagery may contribute to music learning and performance is the Audiation Practice Tool (APT) developed by Susan Williams (Williams, 2019; Williams et al., submitted), who is a baroque trumpet soloist and music pedagogue. In the APT, elements from Gordon's (Gordon, 1979) concept of "audiation," roughly defined as the capacity to create an integrated, multimodal, musically meaningful image, are used to create targets of musical intention (or ideal performance output). This tool, designed to facilitate focusing on the musical meaning of a phrase through multimodal exploration using gestures and visual images rather than the mechanics of playing (i.e., the fingers or the mouth), is based on evidence from motor learning research, implying that more so-called distal focus (task-related but centred on outcome rather than the means of getting there) is helpful when acquiring new skills (Wulf & Prinz, 2001). The APT facilitates this distal focus by directing attention to the musical outcomes but also invites students or performers to reconsider the musical material under consideration by exploring its meaning or feeling, and practicing it in different ways, thereby increasing the time spent thinking about the goal sound as opposed to thinking about manipulating the instruments, or difficult sections coming up. Naturalistic evaluations of APT showed measurable increases in practice success, as well as more reported enjoyment during the practice process (Williams, 2019; Williams et al., submitted).

Although rich mental images conveying meaning are regularly used when instructing or communicating about music performance (as in playing something "with fire"; Black, this volume), and there is previous work on the use of these metaphors or analogies in music pedagogy (Barten, 1998; Woody, 2002), there is very little understanding of how these images are integrated into motor performance. Preliminary questionnaire data from musicians, dancers, and athletes suggest that metaphor use is widespread among all expert groups, and is used not only for different goals, from learning technique to improving performance but also to serve as a communication and collaboration aid, or control nerves (Pietroniro et al., 2016). This further supports the potential usefulness of imagery techniques to help people acquire complex skills in multiple movement expertise domains. However, given the complex nature of these processes, and especially also the high degree of personalization of the used metaphors presupposed to be necessary to support learning and performance, it has also been argued that this domain may not be suitable for more quantitative research (Juslin et al., 2006). However, novel methods that can extract meaningful features from large motion capture datasets may be able to quantify the effects, and provide experimental data that could provide more support for these ideas.

These different ways of applying imagery aspects in music pedagogy and performance all make use of very rich, detailed, and multimodal images, and range from having a mental representation of a musical piece with all its inherent associations, to imagining scenes, emotions, or agents that somehow represent an element of the musical meaning. Clearly, different pedagogical goals will require different approaches, whether it is memorization of complex material, communication or expressiveness, reducing cognitive load during performance, or finding the most effective focus that works for an individual student or performer. More knowledge on the generalizability of these different approaches or techniques will allow further development and practical steps to increase the usability and awareness of imagery as a pedagogical tool.

Discussion

The concepts and studies discussed earlier indicate that various imagery techniques can be suitable tools to support movement acquisition goals, in rehabilitation settings, or aspects of music pedagogy. Some common themes emerge, namely how music imagery, showing additional SMA activation in the brain as compared to music perception as well as when compared to moving without imagery, might contribute something especially useful, as already suggested for pedagogy settings by Halpern and Overy (2019). Additionally, there are some findings suggesting that multimodal imagery through external focus, as applied in the APT (Williams, 2019), is also useful in movement rehabilitation (Kal et al., 2015), further underlining the parallels between both domains.

In rehabilitation specifically, the potential benefits are not only practical, as imagery is cheap and portable, but also potentially cognitive, by enhancing the awareness or experience of the rhythmic cue by the effort needed to generate an image, rather than passively receiving the stimulus. Similarly, in pedagogical settings, the cognitive aspects of practice are intensified, while the strain on the body is arguably reduced (which is a concern for musicians who spend many hours practicing). However, further experimental evaluation of this hypothesized advantage of additional cognitive engagement is needed in order to support this notion. Nonetheless, robust experimental findings underline the benefits of multimodality on motor learning, as well as the involvement of imagery functions in perceptual embedding of movement (Brown & Palmer, 2012, 2013)

To more fully understand the potential and generalizability of these techniques, it is important to recognize that there are individual differences in imagery abilities (cf. Floridou et al., 2022; Isaac & Marks, 1994; Kosslyn et al., 1984) that complicate the extent to which these techniques may work for everyone. In a recent study on imagery abilities over the life span, it was found that even though imagery is thought to interact with cognitive abilities such as working memory and other executive functions, which are known to decline with age, imagery ability does not actually appear to decrease with age, at least up to 65 years; if anything, a mild increase in self-reported vividness of auditory imagery was seen with age (Floridou et al., 2022). Imagery abilities are known to increase with domain-specific training, for instance in music or sports (cf. Herholz et al., 2008; Janata & Paroo, 2006; Ozel et al., 2004), but the potential of training music imagery abilities is yet unclear. In music pedagogy, as described, some individual pedagogues make deliberate use of imagery functions, but imagery is not often explicitly trained in music lessons (although a majority of students believe it is a crucial aspect of musicianship, see Davidson-Kelly et al., 2012). Relatedly, it is not always clear how best to instruct imagery successfully. In rehabilitation settings using movement imagery this issue was already raised as a complicating factor in clinical practice (Malouin & Richards, 2010), and although the instruction "move to the music in your head" appears relatively intuitive, this assumption needs validation.

Taken together, music imagery abilities may be useful for movement recovery, and more generalized imagery techniques may support musical skill acquisition and performance. While many of these concepts still need scientific support to be further developed into broadly applicable protocols and teaching methods, the observation that many clinicians and teachers are already intuitively using these approaches is striking. Future research focusing on refining our understanding of the different types of imagery, the range of individual differences in imagery abilities, and the predictors of these differences, offers to yield a wide variety of applications in both movement rehabilitation and music pedagogy.

Note

1 Only a small group of people are not able to mentally generate visual images, thought to generalize to simulating other sensory modalities—a condition termed congenital

aphantasia (Zeman et al., 2015). It is currently characterized by a score on the VVIQ (Marks, 1973) of ≤ 30 (Wicken et al., 2019; Zeman et al., 2015), its prevalence estimated to lie around 2.4% of the general population (Faw, 2009), also reflected more recently in Floridou et al. (2022).

References

Adolphe, M. B. (2013). *The mind's ear: Exercises for improving the musical imagination for performers, composers, and listeners*. Oxford: Oxford University Press.

Baddeley, A. D., & Andrade, J. (2000). Working memory and the vividness of imagery. *Journal of Experimental Psychology General*, *129*(1), 126–145. https://doi.org/10.1037/0096-3445.129.1.126

Barten, S. S. (1998). Speaking of music: The use of motor-affective metaphors in music instruction. *Journal of Aesthetic Education*, *32*(2), 89–97. https://doi.org/10.2307/3333561

Benoit, C.-E., Dalla Bella, S., Farrugia, N., Obrig, H., Mainka, S., & Kotz, S. A. (2014). Musically cued gait-training improves both perceptual and motor timing in Parkinson's disease. *Frontiers in Human Neuroscience*, *8*, 494. https://doi.org/10.3389/fnhum.2014.00494

Blaisdell, A. P. (2019). Mental imagery in animals: Learning, memory, and decision-making in the face of missing information. *Learning & Behavior*, *47*(3), 193–216. https://doi.org/10.3758/s13420-019-00386-5

Brown, R. M., & Palmer, C. (2012). Auditory-motor learning influences auditory memory for music. *Memory & Cognition*, *40*(4), 567–578. https://doi.org/10.3758/s13421-011-0177-x

Brown, R. M., & Palmer, C. (2013). Auditory and motor imagery modulate learning in music performance. *Frontiers in Human Neuroscience*, *7*, 320. https://doi.org/10.3389/fnhum.2013.00320

Campbell, S. M., & Margulis, E. H. (2015). Catching an earworm through movement. *Journal of New Music Research*, *44*(4), 347–358. https://doi.org/10.1080/09298215.2015.1084331

Davidson-Kelly, K., Moran, N., & Overy, K. (2012, July 23–28). *Learning and memorisation amongst advanced piano students: A questionnaire survey* [Poster presentation]. International Conference on Music Perception and Cognition, Thessaloniki, Greece.

Davidson-Kelly, K., Schaefer, R. S., Moran, N., & Overy, K. (2015). "Total inner memory": Deliberate uses of multimodal musical imagery during performance preparation. *Psychomusicology: Music, Mind, and Brain*, *25*(1), 83–92. https://doi.org/10.1037/pmu000009

Faw, B. (2009). Conflicting intuitions may be based on differing abilities: Evidence from mental imaging research. *Journal of Consciousness Studies*, *16*(4), 45–68.

Floridou, G. A., den Hartog, A., Coers, M., & Schaefer, R. S. (2018, July 23–28). *Neural correlates of movement cued with heard and imagined music* [Poster presentation]. International Conference for Music Perception and Cognition, Graz, Austria.

Floridou, G. A., Peerdeman, K. J., & Schaefer, R. S. (2022). Individual differences in mental imagery in different modalities and levels of intentionality. *Memory & Cognition*, *50*, 29–44. https://doi.org/10.3758/s13421-021-01209-7

Floridou, G. A., Williamson, V. J., Stewart, L., & Müllensiefen, D. (2015). The Involuntary Musical Imagery Scale (IMIS). *Psychomusicology: Music, Mind, and Brain*, *25*(1), 28–36. https://doi.org/10.1037/pmu0000067

Gordon, E. E. (1965). *Musical aptitude profile*. Austin: Houghton Mifflin.

Gordon, E. E. (1979). Developmental music aptitude as measured by the primary measures of music audiation. *Psychology of Music*, *7*(1), 42–49. https://doi.org/10.1177/030573567971005

Halpern, A. R., & Overy, K. (2019). Voluntary auditory imagery and music pedagogy. In M. Grimshaw-Aagaard, M. Walther-Hansen, & M. Knakkergaard (Eds.), *The Oxford handbook of sound and imagination* (Vol. 1, pp. 391–407). Oxford: Oxford University Press.

Halpern, A. R., & Zatorre, R. J. (1999). When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. *Cerebral Cortex*, *9*(7), 697–704. https://doi.org/10.1093/cercor/9.7.697

Halpern, A. R., Zatorre, R. J., Bouffard, M., & Johnson, J. A. (2004). Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia*, *42*(9), 1281–1292. https://doi.org/10.1016/j.neuropsychologia.2003.12.017

Hargreaves, D. J. (2012). Musical imagination: Perception and production, beauty and creativity. *Psychology of Music*, *40*(5), 539–557. https://doi.org/10.1177/0305735612444893

Harrison, E. C., Horin, A. P., & Earhart, G. M. (2019). Mental singing reduces gait variability more than music listening for healthy older adults and people with Parkinson disease. *Journal of Neurologic Physical Therapy*, *43*(4), 204–211. https://doi.org/10.1097/NPT.0000000000000288

Harrison, E. C., McNeely, M. E., & Earhart, G. M. (2017). The feasibility of singing to improve gait in Parkinson disease. *Gait & Posture*, *53*, 224–229. https://doi.org/10.1016/j.gaitpost.2017.02.008

Herholz, S. C., Halpern, A. R., & Zatorre, R. J. (2012). Neuronal correlates of perception, imagery, and memory for familiar tunes. *Journal of Cognitive Neuroscience*, *24*(6), 1382–1397. https://doi.org/10.1162/jocn_a_00216

Herholz, S. C., Lappe, C., Knief, A., & Pantev, C. (2008). Neural basis of music imagery and the effect of musical expertise. *European Journal of Neuroscience*, *28*(11), 2352–2360. https://doi.org/10.1111/j.1460-9568.2008.06515.x

Isaac, A. R., & Marks, D. F. (1994). Individual differences in mental imagery experience: Developmental changes and specialization. *British Journal of Psychology*, *85*(4), 479–500. https://doi.org/10.1111/j.2044-8295.1994.tb02536.x

Janata, P., & Paroo, K. (2006). Acuity of auditory images in pitch and time. *Percept Psychophysics*, *68*(5), 829–844. https://doi.org/10.3758/BF03193705

Juslin, P. N., Karlsson, J., Lindstro, E., Friberg, A., & Schoonderwaldt, E. (2006). Play it again with feeling: Computer feedback in musical communication of emotions. *Journal of Experimental Psychology: Applied*, *12*(2), 79–95. https://doi.org/10.1037/1076-898X.12.2.79

Kal, E. C., van der Kamp, J., Houdijk, H., Groet, E., van Bennekom, C. A. M., & Scherder, E. J. A. (2015). Stay focused! The effects of internal and external focus of attention on movement automaticity in patients with stroke. *PLos One*, *10*(8), e0136917. https://doi.org/10.1371/journal.pone.0136917

Keller, P. E. (2008). Joint action in music performance. In F. Morganti, A. Carassa, & G. Riva (Eds.), *Enacting intersubjectivity: A cognitive and social perspective on the study of interactions* (pp. 205–221). Amsterdam: IOS Press.

Keller, P. E. (2012). Mental imagery in music performance: Underlying mechanisms and potential benefits. *Annals of the New York Academy of Sciences*, *1252*(1), 206–213. https://doi.org/10.1111/j.1749-6632.2011.06439.x

Kleim, J. A., & Jones, T. A. (2008). Principles of experience-dependent neural plasticity: Implications for rehabilitation after brain damage. *Journal of Speech, Language, and Hearing Research*, *51*(1), S225–S239. https://doi.org/10.1044/1092-4388(2008/018)

Korman, M., Raz, N., Flash, T., & Karni, A. (2003). Multiple shifts in the representation of a motor sequence during the acquisition of skilled performance. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(21), 12492–12497. https://doi.org/10.1073/pnas.2035019100

Kosslyn, S. M., Brunn, J., Cave, K. R., & Wallach, R. W. (1984). Individual differences in mental imagery ability: A computational analysis. *Cognition*, *18*(1–3), 195–243. https://doi.org/10.1016/0010-0277(84)90025-8

Maier, M., Ballester, B. R., & Verschure, P. (2019). Principles of neurorehabilitation after stroke based on motor learning and brain plasticity mechanisms. *Frontiers in Systems Neuroscience*, *13*, 74. https://doi.org/10.3389/fnsys.2019.00074

Malouin, F., & Richards, C. L. (2010). Mental practice for relearning locomotor skills. *Physical Therapy*, *90*(2), 240–251. https://doi.org/10.2522/ptj.20090029

Margulis, E. H. (2013). Repetition and emotive communication in music versus speech. *Frontiers in Psychology*, *4*, 167. https://doi.org/10.3389/fpsyg.2013.00167

Marks, D. F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology*, *64*(1), 17–24. https://doi.org/10.1111/j.2044-8295.1973.tb01322.x

Moore, E., Schaefer, R. S., Bastin, M. E., Roberts, N., & Overy, K. (2017). Diffusion tensor MRI tractography reveals increased Fractional Anisotropy (FA) in arcuate fasciculus following music-cued motor training. *Brain and Cognition*, *116*, 40–46. https://doi.org/10.1016/j.bandc.2017.05.001

Moumdjian, L., Moens, B., Maes, P.-J., Van Geel, F., Ilsbroukx, S., Borgers, S., . . . Feys, P. (2019). Continuous 12 minute walking to music, metronomes and in silence: Auditory-motor coupling and its effects on perceived fatigue, motivation and gait in persons with multiple sclerosis. *Multiple Sclerosis and Related Disorders*, *35*, 92–99. https://doi.org/10.1016/j.msard.2019.07.014

Navarro Cebrian, A., & Janata, P. (2010). Influences of multiple memory systems on auditory mental image acuity. *Journal of the Acoustical Society of America*, *127*(5), 3189–3202. https://doi.org/10.1121/1.3372729

Nombela, C., Hughes, L. E., Owen, A. M., & Grahn, J. A. (2013). Into the groove: Can rhythm influence Parkinson's disease? *Neuroscience and Biobehavioral Reviews*, *37*(10), 2564–2570. https://doi.org/10.1016/j.neubiorev.2013.08.003

Ozel, S., Larue, J., & Molinaro, C. (2004). Relation between sport and spatial imagery: Comparison of three groups of participants. *The Journal of Psychology*, *138*(1), 49–63. https://doi.org/10.3200/JRLP.138.1.49-64

Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review*, *76*(3), 241. https://psycnet.apa.org/doi/10.1037/h0027272

Pietroniro, J. D., de Bruin, J. J., & Schaefer, R. S. (2016, July 5–9). *Metaphor use in skilled movement: Music performance, dance and athletics* [Poster presentation]. International Conference for Music Perception and Cognition, San Francisco, CA.

Sampaio-Baptista, C., Sanders, Z. B., & Johansen-Berg, H. (2018). Structural plasticity in adulthood with motor learning and stroke rehabilitation. *Annual Review of Neuroscience*, *41*, 25–40. https://doi.org/10.1146/annurev-neuro-080317-062015

Satoh, M., & Kuzuhara, S. (2008). Training in mental singing while walking improves gait disturbance in Parkinson's disease patients. *European Neurology*, *60*(5), 237–243. https://doi.org/10.1159/000151699

Schaefer, R. S. (2014a). Auditory rhythmic cueing in movement rehabilitation: Findings and possible mechanisms. *Philosophical Transactions of the Royal Society of London: Series B, Biological Sciences*, *369*(1658), 20130402. https://doi.org/10.1098/rstb.2013.0402

Schaefer, R. S. (2014b). Mental representations in musical processing and their role in action-perception loops. *Empirical Musicology Review*, *9*(3), 161–176. http://dx.doi.org/10.18061/emr.v9i3-4.4291

Schaefer, R. S. (2018). Imagery and memory. In R. Ashley & R. Timmers (Eds.), *Routledge companion to music cognition*. London: Routledge.

Schaefer, R. S., Morcom, A. M., Roberts, N., & Overy, K. (2014). Moving to music: Effects of heard and imagined musical cues on movement-related brain activity. *Frontiers in Human Neuroscience*, *8*, 774. https://doi.org/10.3389/fnhum.2014.00774

Schaefer, R. S., & Overy, K. (2015). Motor responses to a steady beat. *Annals of the New York Academy of Sciences*, *1337*(1), 40–44. https://doi.org/10.1111/nyas.12717

Schauer, M., & Mauritz, K. (2003). Musical Motor Feedback (MMF) in walking hemiparetic stroke patients: Randomized trials of gait improvement. *Clinical Rehabilitation*, *17*(7), 713–722. https://doi.org/10.1191/0269215503cr668oa

Seashore, C. E. (1915). The measurement of musical talent. *The Musical Quarterly*, *1*(1), 129–148.

Tartaglia, E. M., Bamert, L., Mast, F. W., & Herzog, M. H. (2009). Human perceptual learning by mental imagery. *Current Biology*, *19*(24), 2081–2085. https://doi.org/10.1016/j.cub.2009.10.060

Thaut, M. H. (2005). *Rhythm, music and the brain*. London: Taylor & Francis.

Thaut, M. H., Mcintosh, G. C., & Hoemberg, V. (2015). Neurobiological foundations of neurologic music therapy: Rhythmic entrainment and the motor system. *Frontiers in Psychology*, *5*, 1185. https://doi.org/10.3389/fpsyg.2014.01185

Wicken, M., Keogh, R., & Pearson, J. (2019). The critical role of mental imagery in human emotion: Insights from Aphantasia. *BioRxiv*, 726844. https://doi.org/10.1101/726844

Williams, S. G. (2019). *Finding focus: Using external focus of attention for practicing and performing music* [Doctoral dissertation, Leiden University]. Research Catalogue for Artistic Research. https://openaccess.leidenuniv.nl/handle/1887/73832.

Williams, S. G., Van Ketel, J. E., & Schaefer, R. S. (submitted). *Practicing musical intention: The effects of external focus of attention on musicians' skill acquisition*.

Woody, R. H. (2002). Emotion, imagery and metaphor in the acquisition of musical performance skill. *Music Education Research*, *4*(2), 213–224. https://doi.org/10.1080/1461380022000011920

Wulf, G., & Prinz, W. (2001). Directing attention to movement effects enhances learning: A review. *Psychonomic Bulletin & Review*, *8*(4), 648–660. https://doi.org/10.3758/BF03196201

Zeman, A., Dewar, M., & Della Sala, S. (2015). Lives without imagery: Congenital aphantasia. *Cortex*, *73*, 378–380. https://doi.org/10.1016/j.cortex.2015.05.019

19 Applied Mental Imagery and Music Performance Anxiety

Katherine K. Finch and Jonathan M. Oakman

Mental imagery refers to a sensory-like experience in the absence of external stimuli and has been described as "seeing with the mind's eye" or "hearing with the mind's ear" (Kosslyn et al., 2001). It can be experienced intentionally (e.g., voluntarily created) or spontaneously (e.g., appears involuntarily) and impacts emotional centres of the brain (e.g., Kim et al., 2007). Because imagery is emotionally impactful and can be brought under intentional control, it has been used to modulate emotions such as anxiety.

Commonly defined as "[t]he experience of marked and persistent anxious apprehension related to musical performance . . . manifested through combinations of affective, cognitive, somatic, and behavioural symptoms" (Kenny, 2010, p. 433), *music performance anxiety* (MPA) is experienced by musicians regardless of their level of expertise. Based on the level of distress and impairment that it causes, MPA can be diagnosed as performance only social anxiety disorder (American Psychiatric Association, 2013). Cognitive theoretical models of social anxiety disorder (Clark & Wells, 1995) suggest that negative spontaneous self-imagery maintains social anxiety. Accordingly, intentional mental imagery manipulation is part of many social anxiety interventions (e.g., McEvoy & Saulsman, 2014) and has also been used to address MPA (e.g., Esplen & Hodnett, 1999). The primary aim of our chapter is to review how existing applied imagery techniques have been used to help musicians manage MPA.

We employ the term *mental preparation imagery* to refer to applied imagery techniques that have been used to help musicians cope with MPA, which we have organized into the following categories: metaphorical imagery, relaxation imagery, systematic desensitization, and mental skills training and imagery. Imagery has been used to help musicians with general experiences of MPA (i.e., not oriented to a particular performance; Stanton, 1994; Whitaker, 1984), in preparation for specific performances (Esplen & Hodnett, 1999; Sisterhen, 2005), or with specific performance-related goals in mind (Cohen & Bodner, 2019; Kageyama, 2007; Osborne et al., 2014). Mental imagery has also been employed to help musicians prepare for specific feared performance situations or experiences (Appel, 1976; Kim, 2005, 2008; Lund, 1972; Nagel et al., 1989; Norton et al., 1978; Reitman, 2001; Rider, 1987; Wardle, 1974), or metaphorically to help musicians prepare for performances (e.g., Stanton, 1994). We further use the term *mental rehearsal* to

DOI: 10.4324/9780429330070-24

denote intentional imagery of specific performance tasks (e.g., shifting or vibrato) that is used to help musicians cope with MPA.

Although the level of procedural detail varies across studies reviewed in our chapter which hampers comparisons, Table 19.1 (available in the Supplementary Materials) provides an overall summary of the studies reviewed later. MPA has been primarily measured through self-report measures, as well as physiological indicators such as heart rate. Additionally, where included in the studies reviewed later, information concerning the sensory modalities, visual imagery perspective,[1] and details regarding how imagery was specifically employed in the interventions is summarized in Table 19.2 (available in the Supplementary Materials).

Metaphorical Imagery

As its name suggests, metaphorical imagery involves imagining symbols rather than imagining the specifics of a performance as in other forms of mental preparation imagery (Black, this volume). However, it can be considered an imagery-based mental preparation technique, as it is designed to help musicians manage general feelings of MPA. Stanton (1994) assigned participants to an intervention condition involving a hypnotic induction and symbolic success imagery designed to reduce MPA, or to a control group. Change in MPA over time was not compared between groups. However, MPA significantly decreased in the intervention condition from pre- to post-test, with decreases maintained at a 6-month follow-up.

Relaxation Imagery

Relaxation imagery has been used as a mental preparation technique by instructing musicians to use performance imagery following a relaxation induction (e.g., progressive muscle relaxation), or by instructing them to imagine performing in a relaxed state. Whitaker (1984) assigned participants to a multicomponent intervention condition involving relaxation imagery or to a control condition. At post-test, MPA was significantly lower in the intervention compared to control group.

Guided relaxation imagery was used as a mental preparation technique to help musicians decrease performance arousal, create feelings of control over performances, and challenge negative appraisals of performing (Esplen & Hodnett, 1999). The authors reported a significant reduction in state anxiety from pre- to post-treatment.

Guided relaxation imagery was also used as mental preparation technique adapted from Suinn's (1994) visuo-motor behaviour rehearsal (VMBR) to reduce MPA (Sisterhen, 2005). Although no significant findings were reported due to a small sample size, MPA decreased for four of the five participants.

Systematic Desensitization

Some mental preparation imagery techniques have incorporated intentional imagery of anxiety-provoking performance situations and experiences. Systematic desensitization involves repeated imaginal exposure to progressively anxiety-provoking stimuli while in a physiologically relaxed state to reduce anxiety. Wolpe (1958) incorporated Jacobson's (1938) progressive muscle relaxation into systematic desensitization to help individuals achieve physiological relaxation while imagining hierarchies of increasingly anxiety-provoking stimuli, with the rationale that repeated exposure to anxiety-provoking stimuli in the presence of an incompatible response (i.e., physiological relaxation), reduces fear. Although contemporary exposure therapy for anxiety disorders is rooted in systematic desensitization (McNally, 2007), theory and exposure research suggest that exposure is effective without relaxation strategies such as progressive muscle relaxation (for a review, see Moscovitch et al., 2009). However, many MPA studies have utilized systematic desensitization—with concurrent progressive muscle relaxation—and are thus reviewed later. The content of imagery used in anxiety hierarchies (e.g., sensory modalities and performance situations) varies in systematic desensitization interventions depending on the specific anxiety hierarchies used by researchers.

Two experiments have compared systematic desensitization with insight-oriented relaxation and a control group (Lund, 1972; Wardle, 1974). Lund (1972) found no differences between groups on performance quality or MPA at post-test. Wardle (1974) found differences between the groups, with MPA decreasing in the systematic desensitization and insight-oriented relaxation conditions.[2] Two further studies have evaluated a combination of systematic desensitization with other interventions. Appel (1976) assigned participants to a systematic desensitization with in vivo desensitization (i.e., live performances), music analysis with performance rehearsal, or a control group. MPA decreased significantly from pre to post-test, with decreases in the systematic desensitization condition greater than those in the music analysis with performance rehearsal and control group. Nagel et al. (1989) assigned participants to a cognitive behavioural therapy intervention or a control group. Although a formal systematic desensitization protocol was not adopted in this study, all of the important components were present (i.e., imaginal exposure to a hierarchy of anxiety-provoking performance situations following progressive muscle relaxation). At endpoint, MPA and trait anxiety were significantly lower in the intervention condition compared to the control condition.

Systematic desensitization has also been incorporated into two case studies using various other treatment techniques, both with positive results (Norton et al., 1978; Rider, 1987).

Music Therapy and Systematic Desensitization

Given the unique relationship that musicians have with music, several researchers have used music therapy techniques to enhance systematic desensitization to reduce MPA. Reitman (2001) assigned participants to a verbal-coping systematic desensitization, a music-assisted coping systematic desensitization, or control condition. Sessions were held in groups and were not led by a certified music therapist. There were no significant differences between the groups on MPA. Kim

(2005, 2008) has tested a Music Therapy Improvisation and Desensitization Protocol for the amelioration of MPA (based on Reitman's MCSD described earlier) in individual sessions led by a certified music therapist. An initial pilot study of six participants yielded encouraging results with trait anxiety decreasing at post-test (Kim, 2005). In a follow-up experiment, Kim (2008) assigned participants to a condition based on the Music Therapy Improvisation and Desensitization Protocol or to a music-assisted progressive muscle relaxation and imagery condition. Change from baseline to endpoint was evident for both groups; however, there were no significant differences in MPA or state anxiety between groups at endpoint.

Mental Skills Training and Imagery

Researchers have noted that interventions which are solely aimed at reducing anxiety do not address the complex cognitive and motor demands of performing a musical instrument under stress (Osborne et al., 2014). Mental skills training interventions draw on sport psychology and integrate numerous strategies to help performers manage anxiety and optimize performance, including mental preparation imagery and mental rehearsal (Osborne et al., 2014). The number of mental skills training interventions for musicians is growing, and mental preparation and mental rehearsal have been used in mental skills training interventions in a variety of ways.

Three studies (Cohen & Bodner, 2019; Kageyama, 2007; Osborne et al., 2014) have drawn on mental preparation and rehearsal techniques based on Greene (2002). One of the techniques from Greene's work that has been used is Centring Down, which can be viewed as a mental preparation technique to help musicians cope with MPA. Centring Down follows a seven-step process including setting a clear goal or intention for a performance and incorporates imagery to see (through first- and third-person visual imagery), hear, and feel oneself play well. Mental skills interventions have also incorporated mental rehearsal based on Greene (2002), who further suggests that mental rehearsal can help with spontaneous negative imagery that musicians might experience (e.g., making mistakes).

Tests of these interventions have yielded mixed results. In a mental skills intervention, Kageyama (2007) examined the effects of an arousal control training, attention control training (incorporating Centring Down), or control group to enhance performance quality and state anxiety was also measured as an MPA proxy. No significant differences between the groups were reported on either state anxiety or performance quality. Osborne et al. (2014) investigated a mental skills intervention based on Greene (2002, 2012, as cited in Osborne et al., 2014) incorporating Centring Down and mental rehearsal, finding decreases on the MPA subscale of their dependent measure. Cohen and Bodner (2019) investigated the helpfulness of a mental skills intervention—incorporating Centring Down and mental rehearsal—designed to reduce MPA and optimize performance and flow-state compared to a control group. There was a significant decrease in MPA in the intervention compared to control condition from pre- to post-test. Only the intervention group engaged in performances as part of the intervention. State anxiety measured before performances significantly decreased while performance quality significantly increased from pre- to post-test.

Clark and Williamon (2011) investigated a mental skills intervention designed to manage MPA, increase self-efficacy, optimize performance, and increase mental skills use compared to a control condition. Mental preparation imagery as well as mental rehearsal were used. The treatment group had superior self-efficacy endpoint scores relative to the control group, but there were no differences in state MPA between the groups. Braden et al. (2015) included a cognitive behavioural intervention incorporating mental skills to help musicians manage MPA and improve their performance quality and a control condition. Post-intervention, MPA was significantly lower in the intervention group compared to control condition.

Integrative Summary and Future Directions

Intentional mental imagery manipulation has been used in different ways to help musicians cope with MPA with varying degrees of empirical support. Several limitations must be taken into account when reviewing existing research including small sample sizes and designs (e.g., one-arm) which limit our understanding of the causal relationships between mental imagery and MPA. Indeed, many of the positive findings reviewed were from uncontrolled or quasi-experimental designs. Additionally, different MPA measures have been used in existing work; consensus in MPA measurement would aid comparisons regarding the helpfulness of imagery across studies. Furthermore, applied imagery models (e.g., Gregg & Clark, 2007; Martin et al., 1999) suggest that imagery ability moderates the efficacy of mental imagery, and future research should include measures of imagery ability (Schaefer, this volume). The field would also benefit from including more information about how imagery has been deployed in interventions. In particular, visual imagery perspective (e.g., first or third person) differentially impacts emotions (Libby & Eibach, 2011), yet few studies have included information about imagery perspective.

Of the studies reviewed, some of the strongest evidence comes from mental skills interventions utilizing mental preparation and rehearsal, although the quasi-experimental designs of these studies must be taken into consideration (Braden et al., 2015; Clark & Williamon, 2011; Cohen & Bodner, 2019; Osborne et al., 2014). Mental rehearsal is positively associated with performance outcomes (Driskell et al., 1994; Kenny, 2006)[3] suggesting that improving performance quality is a way that performers can enhance their confidence for future performances. The field would benefit from investigating whether mental rehearsal—used on its own—has a downstream effect on MPA (e.g., perhaps via inhibition of anxiety-laden spontaneous images or simply through increased confidence in performance).

With the exception of metaphorical imagery, the mental preparation techniques reviewed can be conceptualized as a form of imaginal exposure, in which musicians are exposed to aspects of anxiety-provoking performance experiences. However, inconsistent findings and methodological limitations preclude a clear understanding of whether certain approaches (e.g., relaxation) are more effective than others. Additionally, recent attempts to enhance mental preparation imagery (e.g., systematic desensitization with music therapy) have not yielded positive findings.

Current best-practice guidelines for exposure suggest that anxiety control strategies should be avoided (Barlow, 2014) to allow for engagement with feared stimuli (Foa et al., 2006; for a review of alternative perspectives, see Blakey & Abramowitz, 2016). Using relaxation imagery or anxiety-imagery while physiologically relaxed may involve anxiety control strategies to suppress the expression or experience of performance anxiety for musicians with MPA. Future imagery-based MPA research would benefit from including imagery strategies used in sport psychology which incorporate high arousal imagery without concurrent relaxation strategies. Such strategies help athletes reinterpret competitive anxiety in a more facilitative direction (Cumming et al., 2007), and are in line with best-practice exposure guidelines.

Greene (2002) suggests that mental rehearsal may limit interference from negative spontaneous imagery related to performing. This is an intriguing suggestion, especially given the role of negative spontaneous imagery in the maintenance of social anxiety (Clark & Wells, 1995). To this end, research suggests that imagery ability protects against interference—presented during encoding or retrieval—during musical task performance (Brown & Palmer, 2013). Although such findings may not generalize to interference from negative spontaneous imagery, future research should investigate whether different forms of intentional mental imagery (e.g., mental preparation) protect against negative spontaneous imagery, and whether this effect is pronounced in musicians with higher imagery ability.

In summary, our review raises important questions regarding how intentional mental imagery manipulation has been used in previous research and highlights areas for future research. In particular, the field would benefit from investigating theory-driven imagery strategies in adequately powered experimental designs to elucidate the causal links between imagery and MPA. Furthermore, such research should also investigate potential mechanisms of the effectiveness of mental imagery (e.g., interference from negative spontaneous imagery). Future research in these areas will provide musicians with a clearer understanding of how to utilize intentional mental imagery manipulation to cope with MPA.

Notes

1 Noting the perspective taken during visual imagery is important as visual imagery perspectives (i.e., first and third person) differentially impact emotion (e.g., Libby & Eibach, 2011).
2 However, simple effects tests were not reported.
3 However, this effect is negatively associated with the degree to which a task requires coordination (Driskell et al., 1994).

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington: American Psychiatric Publishing.

Appel, S. S. (1974). *Modifying solo performance anxiety in adult pianists* (Order No. 7426580). Available from ProQuest Dissertations & Theses Global. (302684108). http://search.proquest.com.proxy.lib.uwaterloo.ca/docview/302684108?accountid=14906

Appel, S. S. (1976). Modifying solo performance anxiety in adult pianists. *Journal of Music Therapy*, *13*(1), 2–16. https://doi-org.proxy.lib.uwaterloo.ca/10.1093/jmt/13.1.2

Barlow, D. H. (Ed.). (2014). *Clinical handbook of psychological disorders: A step-by-step treatment manual* (5th ed.). New York, NY: The Guilford Press.

Blakey, S. M., & Abramowitz, J. S. (2016). The effects of safety behaviors during exposure therapy for anxiety: Critical analysis from an inhibitory learning perspective. *Clinical Psychology Review*, *49*, 1–15. https://doi.org/10.1016/j.cpr.2016.07.002

Borkovec, T. D. (1976). Physiological and cognitive processes in the regulation of anxiety. In G. E. Schwartz & D. Shapiro (Eds.), *Consciousness and self-regulation* (pp. 261–312). Boston, MA: Springer.

Braden, A. M., Osborne, M. S., & Wilson, S. J. (2015). Psychological intervention reduces self-reported performance anxiety in high school music students. *Frontiers in Psychology*, *6*, 195. https://doi.org/10.3389/fpsyg.2015.00195

Brown, R. M., & Palmer, C. (2013). Auditory and motor imagery modulate learning in music performance. *Frontiers in Human Neuroscience*, *7*, 320. https://doi.org/10.3389/fnhum.2013.00320

Clark, D. M., & Wells, A. (1995). A cognitive model of social phobia. In R. G. Heimberg, M. R. Liebowitz, D. A. Hope, & F. R. Schneier (Eds.), *Social phobia: Diagnosis, assessment and treatment* (pp. 69–93). New York: Guilford Press.

Clark, T., & Williamon, A. (2011). Evaluation of a mental skills training program for musicians. *Journal of Applied Sport Psychology*, *23*(3), 342–359. https://doi.org/10.1080/10413200.2011.574676

Cohen, S., & Bodner, E. (2019). Music performance skills: A two-pronged approach: Facilitating optimal music performance and reducing music performance anxiety. *Psychology of Music*, *47*(4), 521–538. https://doi.org/10.1177/0305735618765349

Cox, R. H., Martens, M. P., & Russell, W. D. (2003). Measuring anxiety in athletics: The revised competitive state anxiety inventory-2. *Journal of Sport and Exercise Psychology*, *25*(4), 519–533. https://doi.org/10.1123/jsep.25.4.519

Cumming, J., Olphin, T., & Law, M. (2007). Self-reported psychological states and physiological responses to different types of motivational general imagery. *Journal of Sport and Exercise Psychology*, *29*(5), 629–644. https://doi.org/10.1123/jsep.29.5.629

Derogatis, L. R. (2001). *Brief Symptom Inventory-18: Administration, scoring, and procedures manual*. Minneapolis, MN: NCS Assessments.

Driskell, J. E., Copper, C., & Moran, A. (1994). Does mental practice enhance performance? *Journal of Applied Psychology*, *79*(4), 481–492. https://doi.org/10.1037/0021-9010.79.4.481

Esplen, M. J., & Hodnett, E. (1999). A pilot study investigating student musicians' experiences of guided imagery as a technique to manage performance anxiety. *Medical Problems of Performing Artists*, *14*, 127–132. www.sciandmed.com/mppa/journalviewer.aspx?issue=1095&article=1049

Foa, E. B., Huppert, J. D., & Cahill, S. P. (2006). Emotional processing theory: An update. In B. O. Rothbaum (Ed.), *Pathological anxiety: Emotional processing in etiology and treatment* (pp. 3–24). New York: Guilford Press.

Greene, D. (2002). *Performance success: Performing your best under pressure*. New York: Routledge.

Gregg, M., & Clark, T. (2007). Theoretical and practical applications of mental imagery. In A. Williamon & D. Coimbra (Eds.), *Proceedings of the international symposium on performance science 2007* (pp. 295–300). Utrecht: European Association of Conservatoires.

Jacobson, E. (1938). *Progressive relaxation: A physiological and clinical investigation of muscular states and their significance in psychological and medical practice*. Chicago: University of Chicago Press.

Kageyama, N. J. (2007). *Attentional focus as a mediator in the anxiety-performance relationship: The enhancement of music performance quality under stress* [Doctoral Dissertation, Indiana University]. ProQuest Dissertations & Theses Global. http://search.proquest.com.proxy.lib.uwaterloo.ca/docview/304849282?accountid=14906

Kenny, D. T. (2006). Music performance anxiety: Origins, phenomenology, assessment and treatment. *Context: Journal of Music Research*, *31*, 51–64.

Kenny, D. T. (2010). The role of negative emotions in performance anxiety. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 425–451). New York: Oxford University Press.

Kim, S. E., Kim, J. W., Kim, J. J., Jeong, B. S., Choi, E. A., Jeong, Y. G., . . . Ki, S. W. (2007). The neural mechanism of imagining facial affective expression. *Brain Research*, *1145*, 128–137. https://doi.org/10.1016/j.brainres.2006.12.048

Kim, Y. (2005). Combined treatment of improvisation and desensitization to alleviate music performance anxiety in female college pianists: A pilot study. *Medical Problems of Performing Artists*, *20*(1), 17–24.

Kim, Y. (2008). The effect of improvisation-assisted desensitization, and music-assisted progressive muscle relaxation and imagery on reducing pianists' music performance anxiety. *Journal of Music Therapy*, *45*(2), 165–191. https://doi-org.proxy.lib.uwaterloo.ca/10.1093/jmt/45.2.165

Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience*, *2*(9), 635–642. https://doi.org/10.1038/35090055

Lehrer, P. M., Goldman, N. S., & Strommen, E. F. (1990). A principal components assessment of performance anxiety among musicians. *Medical Problems of Performing Artists*, *5*(1), 12–18.

Libby, L. K., & Eibach, R. P. (2011). Visual perspective in mental imagery: A representational tool that functions in judgment, emotion, and self-insight. In J. M. Olson & M. P. Zanna (Eds.), *Advances in experimental social psychology* (Vol. 44, pp. 185–245). Amsterdam: Elsevier. https://doi.org/10.1016/B978-0-12-385522-0.00004-4

Lund, D. R. (1972). *A comparative study of three therapeutic techniques in the modification of anxiety behavior in instrumental music performance* [Doctoral dissertation, The University of Utah]. ProQuest Dissertations & Theses Global. http://search.proquest.com.proxy.lib.uwaterloo.ca/docview/302733226?accountid=14906

Martin, K., Moritz, S., & Hall, C. (1999). Imagery use in sport: A literature review and applied model. *The Sport Psychologist*, *13*(3), 245–268. https://doi.org/10.1123/tsp.13.3.245

McEvoy, P. M., & Saulsman, L. M. (2014). Imagery-enhanced cognitive behavioural group therapy for social anxiety disorder: A pilot study. *Behaviour Research and Therapy*, *55*, 1–6. https://doi.org/10.1016/j.brat.2014.01.006

McNally, R. J. (2007). Mechanisms of exposure therapy: How neuroscience can improve psychological treatments for anxiety disorders. *Clinical Psychology Review*, *27*(6), 750–759. https://doi.org/10.1016/j.cpr.2007.01.003

Moscovitch, D. A., Antony, M. M., & Swinson, R. P. (2009). Exposure-based treatments for anxiety disorders: Theory and process. In M. M. Antony & M. B. Stein (Eds.), *Oxford library of psychology: Oxford handbook of anxiety and related disorders* (pp. 461–475). New York: Oxford University Press.

Nagel, J. J., Himle, D. P., & Papsdorf, J. D. (1989). Cognitive-behavioural treatment of musical performance anxiety. *Psychology of Music*, *17*(1), 12–21. https://doi.org/10.1177%2F0305735689171002

Norton, G. R., MacLean, L., & Wachna, E. (1978). The use of cognitive desensitization and self-directed mastery training for treating stage fright. *Cognitive Therapy and Research*, *2*(1), 61–64. https://doi.org/10.1007/BF01172513

Osborne, M. S., Greene, D. J., & Immel, D. T. (2014). Managing performance anxiety and improving mental skills in conservatoire students through performance psychology training: A pilot study. *Psychology of Well-Being*, *4*(1), 18. https://doi.org/10.1186/s13612-014-0018-3

Osborne, M. S., & Kenny, D. T. (2005). Development and validation of a music performance anxiety inventory for gifted adolescent musicians. *Journal of Anxiety Disorders*, *19*(7), 725–751. https://doi.org/10.1016/j.janxdis.2004.09.002

Reitman, A. D. (2001). The effects of music-assisted coping systematic desensitisation on music performance anxiety. *Medical Problems of Performing Artists*, *16*(3), 115–125. https://doi.org/10.21091/mppa.2001.3020

Rider, M. S. (1987). Music therapy: Therapy for debilitated musicians. *Music Therapy Perspectives*, *4*(1), 40–43. https://doi.org/10.1093/mtp/4.1.40

Ritchie, L., & Williamon, A. (2011). Measuring distinct types of musical self-efficacy. *Psychology of Music*, *39*, 328–344. https://doi.org/10.1177/0305735610374895

Sheehan, P. W. (1967). A shortened form of Betts' questionnaire upon mental imagery. *Journal of Clinical Psychology*, *23*(3), 386–389. https://doi.org/10.1002/1097-4679(196707)23:3<386::AID-JCLP2270230328>3.0.CO;2-S

Shorkey, C. T., & Whiteman, V. L. (1977). Development of the rational behavior inventory: Initial validity and reliability. *Educational and Psychological Measurement*, *37*(2), 527–534.

Sisterhen, L. A. (2005). *The use of imagery, mental practice, and relaxation techniques for musical performance enhancement* [Doctoral Dissertation, The University of Oklahoma]. ProQuest Dissertations & Theses Global. http://search.proquest.com.proxy.lib.uwaterloo.ca/docview/276093741?accountid=14906

Spielberger, C. D. (1980). *Test anxiety inventory*. Palo Alto, CA: Consulting Psychologists Press.

Spielberger, C. D. (1989). Stress and anxiety in sports. In D. Hackfort & C. D. Spielberger (Eds.), *Anxiety in sports: Theory and assessment* (pp. 3–17). New York, NY: Hemisphere Publishing Corporation.

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the state: Trait anxiety inventory, form Y, self-evaluation questionnaire*. Palo Alto, CA: Consulting Psychologists Press Inc.

Stanton, H. E. (1994). Reduction of performance anxiety in music students. *Australian Psychologist*, *29*(2), 124–127. https://doi.org/10.1080/00050069408257335

Suinn, R. M. (1994). Visualization in sports. In A. A. Sheikh & E. R. Korn (Eds.), *Imagery in sports and physical performance* (pp. 23–42). Amityville: Baywood Publishing Company, Inc.

Wardle, A. (1974). Behavior modification by reciprocal inhibition of instrumental music performance anxiety. *Journal of Band Research*, *11*(1), 18. http://search.proquest.com.proxy.lib.uwaterloo.ca/docview/1312125909?accountid=14906

Watkins, J. G., & Farnum, S. E. (1954). *The Watkins-Farnum Performance Scale*. Winona, MN: Hal Leonard Music, Incorporated.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measure of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070. http://dx.doi.org/10.1037/0022-3514.54.6.1063

Whitaker, C. S. (1984). *The modification of psychophysiological responses to stress in piano performance (pedagogy, biofeedback*) [Doctoral Dissertation, Texas Tech University]. ProQuest Dissertations & Theses Global. http://search.proquest.com.proxy.lib.uwaterloo.ca/docview/303323997?accountid=14906

Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition*. Palo Alto: Stanford University Press.

Zimmerman, B., & Martinez-Pons, M. (1988). Construct validation of a strategy model of student self-regulated learning. *Journal of Educational Psychology*, *80*, 284–290. http://dx.doi.org/10.1037/0022-0663.80.3.284

20 In Search of a Story

Guided Imagery and Music Therapy

Helena Dukić

Musicians and music lovers have always been curious about the true nature and function of music. What is its purpose? Does music need to relate to extra-musical stories and visual imagery (people, events, or environments) to be meaningful? A variety of different interpretations of musical meaning emerged attempting to integrate the emotional experience with the semantic meaning. Several researchers have proposed different theories to explain the meaning of music (Margulis, 2017; Maus, 1988; Robinson & Hatten, 2012). However, all of them centred around the notion of music as a narrative form, that is, a sequence of meaningful events that acquire the following form: setup-confrontation-resolution (Traupman, 1966). The impulse to narrativize has been characterized as distinctively human, shaping the way people make sense of their lives (McAdams, 2006), so it is natural to presume that the same principle is used to understand music.

A music discipline highlighting the relation between music and narrative is Guided Imagery and Music (GIM), a form of music therapy (see also Fachner, this volume). Developed by Helen Bonny in the 1970s, GIM uses the connection between music pieces and clients' personal narratives for therapeutic purposes by allowing the clients to explore their life events while being contained and guided by music. In GIM, clients listen to a series of music pieces in a state of deep relaxation and as a result, develop spontaneous imagery that is related to the music's temporal structure (Bonny, 2010). The method is called "guided," because the clients are being guided by the music they listen to; music guides their emotional and imaging experience and offers a "container" for the whole therapeutic experience. Bonny proposed that music structure "contained" an emotional experience of the client and provided a "vital centre for personal grounding" (Bonny, 1989, p. 7). For example, if a client is overwhelmed by grief, music played should have a small container (a predictable melody, supportive harmonic and rhythmic structure, and little fluctuation of dynamics). *Imagery* is defined as dream-like imagery that, unlike dreams, can be recounted and remembered after the session (Bonny, 2010). The types of imagery emerging vary and can include descriptions of various settings, different characters, and different physical and mental situations in which the protagonist of the story engages, and a wide array of emotional states. Thus, forms of imagery include not only visuals but also olfactory, tactile, and kinaesthetic imagery, as well as memories and feelings (Bonny, 2010).

DOI: 10.4324/9780429330070-25

The music pieces that the clients listen to during sessions are put together into programmes that bear programmatic names (Relationships, Death-Rebirth, Nurturing, etc.). A GIM therapist selects a programme, depending on the clients' emotional state and the aim of the particular therapeutic session. Music used in GIM sessions is primarily Western classical music, and each programme can last between 25 and 45 minutes. Music selections chosen contribute to the exploration of levels of consciousness not usually available to normal awareness and thus offer the client access to repressed emotions. This is achieved by conducting a deep relaxation exercise with the client in order to initiate an Altered State of Consciousness before the music is played. The induction has two elements: physical relaxation and psychological concentration. Relaxation of the musculature is achieved by a number of different relaxation techniques by verbal suggestion, most complete of which is Progressive Relaxation (Jacobson, 1938). Concentration is achieved by focusing client's attention on one internal stimulus (an imagined place such as a meadow or a beach) to the exclusion of all other stimuli. After the relaxation and the concentration which happen simultaneously, a selected music programme is played. During the listening, the therapist asks the client questions about emerging imagery, deepening their experience and the Altered State of Consciousness by encouraging the client to verbally relate their impressions, imagery, and feelings as they occur in response to music. Once the music is finished, the therapist returns the client to a normal state of consciousness and integrates the experience by offering the client to draw a mandala portraying their imagery or by talking to them about the imagery that occurred. No interpretation of imagery is given by the therapist during or after the session; therapeutic progress proceeds through music as a catalyst and through allowing and persuasive attitude of the therapist who gently pushes the client deeper into their imaging experience by being supportive and understanding of any imagery and feelings that may occur. The session usually starts and ends with a brief conversation aimed at identifying the main issue the client experiences and at integrating the listening experience.

Imagery is the crucial part of every GIM session. Imagery in various modalities is understood as a direct expression of the conscious and unconscious psyche, client's resources and their difficulties as they are shaped by important relationships and events. The imagery presents as symbols that represent unprocessed bodily, sensory, and emotional experiences. Imagery can encompass various combinations of archetypal, psychodynamic, aesthetic, and transpersonal imagery. Clients are encouraged to interact with the imagery in order to gain a deeper understanding of the imagery. Furthermore, clients are encouraged to understand the meaning of the mental image within their own framework, without the external interpretational system. For example, an image of a tree can mean very different things for different clients: A client might be one with the tree and its branch movements; a tree might symbolize nurturing, protection, and growth; the tree may represent relationships and self-image.

GIM offers a framework for empirical research on the relation between music and emerging imagery as well as how music affects the non-auditory experiences such as emotions and visual processing. The limited research that exists has focused on the quality (vividness) of imagery emerging (Grocke, 1999; McKinney, 1990) and the content of the imagery (such as characters and symbols) but did not relate it to a narrative form (Lewis, 1999; Wilber, 1993). Finally, Bruscia and his colleagues conducted a study where he observed that there are certain consistencies in the imagery of different clients, meaning that most clients listening to the same GIM programme experienced similar themes in their imagery (Bruscia et al., 2005). This opened questions about the semiotic meaning of music and strongly suggested a vital role GIM might play in the relationship between music and narrative, providing a great ground for exploring this symbiosis.

First, however, we must define the structural components of a narrative. Aristotle defined a narrative form as a three-part plot: the beginning, the middle, and the end. In 19th century, Freytag (1900) proposed a five-part structure of a narrative plot: exposition, rising action, climax, falling action, and dénouement. His plot structure was later expanded by Joseph Campbell who created a notion of a "monomyth" (Campbell, 1949), a skeleton of a narrative that is similar in all types of stories. Finally, David Herman suggests that a prototypical narrative consists of settings, protagonist's feelings and actions and other characters that create events (Herman, 2009). I thus created a summary of a narrative definition: A narrative is a meaningful set of events that have basic elements (a setting, characters, a plot, and a meaning) which appear in the following three-part form: setup, confrontation, and resolution. A setup is characterized by passive behaviour of the protagonist who mainly observes the environment, confrontation by their active behaviour and physical and emotional involvement with their surroundings and resolution again by the protagonist's passivity. Thus, we could say that the blueprint of a three-part narrative form can be summarized as follows: passive–active–passive.

Defining the narrative elements in a literary work presented a challenge to narrative theorists. A direct connection between characters or settings and musical works raised even bigger issues. Di Bona has commented on the existence of the so-called metaphorical space and time that music evoked, referring to spatial imagery and metaphors which are not necessarily related to the spatial properties of sound sources (Di Bona, 2017). Music can also suggest emotional content to which the listeners apply metaphorical concepts of a certain character. Robinson and Hatten's (2012) suggestions of the existence of a virtual "person" in the music pieces with which the listeners empathize imply that music is capable of portraying a character or a person by expressing a mixture of feelings that change and develop over time.

In a study I conducted with my colleagues (Dukić et al., 2021), we used the GIM method to evaluate which types of characters and settings are elicited by music in the form of imagery in GIM sessions as compared to the literary narrative forms in fairy tales. In the study, we categorized the imagery experienced using narrative elements (Herman et al., 2012) and compared them to the same narrative elements found in fairy-tales (Settings: Flora, Fauna, Events, Structures; Protagonist: Feelings, Actions; and Characters). Fairy-tales were used as a control group, because they were standard narratives that did not have a common theme or moral and therefore represented a neutral material. GIM programmes called "Nurturing" and "Death-Rebirth" were used in the sessions to elicit the imagery. Results show that narrative elements of Structures, Flora, Fauna, and Feelings occurred significantly more often in participants' GIM imagery than in the fairy-tale controls in both GIM programmes. These categories are plot-static: they do not generate active relationships between characters. Events, Actions and Characters occurred significantly less often in GIM imagery. This marks the first and major difference between music and a literally narrative form; music does not seem to suggest imagery that serves the function of plot development in the same quantity the literary narratives do.

Besides a setting and the characters, an important narrative element is also its plot. Lerdahl and Jackendoff (1983) argued that the tension-relaxation exchange conveyed by the music features was the basic force of temporal progression of music, alluding to a set of emotional expectation and resolutions that music creates using harmony and melody. Nattiez (2013) also acknowledged the pattern of tensions and relaxations in music, using the term "proto-narrative" to indicate the fact that music imitates the real narrative and is thus only capable of expressing the blueprint of a story. This study is innovative in its goal of quantitatively examining whether a piece of music can elicit visual imagery that takes on a three-part narrative form throughout the musical piece. Since the three-part structure consisting of a setup, confrontation, and resolution is widely accepted as a standard model of a narrative form (Field, 2007; Herman, 2009), this model will be used as a reference in the current study. If music is capable of prompting the exchange of active and passive imagery, creating a blueprint of a three-part narrative form (passive–active–passive), then we could say that music has a plot.

The Current Study

GIM provides a unique foundation for the current study by granting an insight into a visual imprint music leaves in the minds of the listeners. By concentrating on the actions of the protagonist of the music-evoked imagery, this study will examine if music is capable of prompting the exchange of active (the protagonist engages in activities) and passive (the protagonist observes) behaviours in a way that creates a blueprint of a three-part narrative form (passive–active–passive). The aim of the study is to evaluate whether music-evoked visual imagery from a GIM programme appears in a three-part narrative form. In Lerdahl and Jackendoff's theory (1983), the more significant the contrast between the two music events, the more the alteration between tensions and relaxations will be noticeable to the listeners. Margulis (2017) also suggests that contrast in successive music features makes listeners more likely to hear music in terms of a narrative. It is thus hypothesized that the imagery of the following pieces of the "Nurturing" programme will assume a three-part narrative form: Britten, *Sentimental Sarabande*; Berlioz, *Shepard's Farewell*; Berlioz, *Flight to Egypt*; Massenet, *Under the Linden Trees*; and Canteloube, *Berceuse*. Imagery in Walton, *Touch Her Soft Lips*, and Puccini, *Humming Chorus*, is not expected to produce a three-part narrative form. The subjective listener response of the author suggested that the mentioned five pieces exhibited a greater contrast between tensions and relaxations than the remaining two, supporting the Lerdahl and Jackendoff's (1983) theory and assuming that those will most likely elicit a narrative form of imagery.

Method

Participants

Twenty-three undergraduate psychology students and volunteers from the local community centres in Zagreb, Croatia (19 female, 4 male, age range 20–89 years, $M = 26.6$, $SD = 3.8$), all non-musicians took part in this study. Twenty of them were first-time GIM participants and three of them were experienced GIM participants (had five or more GIM sessions). All the participants spoke Croatian throughout the sessions.

Stimuli

The music used in the study was Bonny's "Nurturing" programme (35 minutes long). The "Nurturing" programme consists of seven music pieces: Britten: *Sentimental Sarabande*, from Simple Symphony no. 4 (6 minutes 41 seconds, instrumental), Walton: *Touch Her Soft Lips* (1 minute 51 seconds, instrumental), Berlioz: *Flight to Egypt*, from The Childhood of Christ, op. 25 (6 minutes 56 seconds, instrumental), Berlioz: *Shepherd's Farewell*, from The Childhood of Christ, op. 25 (5 minutes, vocal), Puccini: *Humming Chorus*, from Madam Butterfly (2 minutes 47 seconds, vocal), Massenet: *Under the Linden Trees*, from Suite Alsatian scenes, no. 7 (4 minutes 59 seconds, instrumental), Canteloube: *Berceuse*, from Songs of Auvergne (3 minutes 13 seconds, vocal).

Procedure

All the participants underwent a standard GIM session, lasting 1 hour 30 minutes. The GIM session with each participant was conducted individually. The therapist was the author of this chapter. Each GIM session had four stages, following the standard formal division: The first stage included an initial interview with each participant asking about their music education, GIM experience, and their age and gender. The definition of imagery was explained to each participant: "Imagery includes any spontaneous visual imagery that occurs during listening." In the second stage, a deep relaxation exercise was conducted lasting 5 minutes. The participants were given the following focus: "Let the music take you where you need to go." In the third stage of the session, the "Nurturing" programme was played. The participants were asked to report any imagery the moment it occurred during the listening and to describe it in as much detail as possible. The therapist sat next to the participant asking non-leading questions such as "What do you see now?" In the fourth stage, when the music stopped, the therapist applied the standard procedure to return participants from the non-ordinary state of consciousness (Bruscia & Grocke, 2002): Participants were encouraged to start moving their limbs, start being aware of the sounds in the room and open their eyes when ready. This was followed by a brief conversation to reflect on the experience and the imagery that emerged. These comments were not included into the imagery analysis, but the participants' comments were available to the coders as an additional explanation of the imagery. An audio recording of all four stages of the sessions was made. Stage three and four were transcribed. The participants' comments from the fourth stage interpreting their imagery were made available to the coders in a written form.

Results

Transcription

Single-track recordings of music and participants' statements were transcribed using MAXQDA12 software for qualitative data analysis (VERBI Software, 2015). Each of the 23 transcripts was divided into paragraphs, each corresponding to approximately 10 seconds of sound with boundaries occurring between sentences. Paragraphs were divided into sentences, each paragraph containing on average one sentence.

Imagery Categorization

The imagery that appeared in the transcriptions of participant's statements during listening was categorized into passive or active imagery types. The qualitative analysis included a distinction between passive and active imagery types and was conducted by three coders (2 females, 1 male, age 19–22 years, M = 20.6, SD = 1.1) who were asked to read every paragraph of the transcribed imagery and decide which type the imagery belonged to. The coders were provided with the following imagery descriptions which they based their decisions on. Passive type: (1) the protagonist of the imagery described the environment around them (nature scenes, cities, animals, or people) and did not engage with it actively, (2) the environment was static. For example: "I'm in a forest and I can see flowers growing next to the path." Active imagery: (1) the protagonist was engaged in either physical or psychological action (2) the environment was active; other characters were animated, structures or nature scenes were active. For example: "I'm climbing up a hill and it's very hard because there is a storm around me." Each coder was provided with 23 transcripts of participants' imagery. Each transcript was divided into paragraphs of one sentence and each coder had to decide whether the sentence in a given paragraph belonged to passive or active imagery type. If a particular imagery type in a given paragraph was selected by two or more coders, it was regarded as the dominant imagery type in that paragraph. The dominant imagery types in all paragraphs of 23 transcripts were counted, their maximum number ranging from 0 (none of the participants had this type of imagery in this paragraph) to 23 (all 23 participants had this type of imagery in this paragraph). The number of participants experiencing certain types of imagery in each paragraph was then shown in a figure for each composition separately in order to see if the distribution of imagery types followed the three-part narrative form (passive–active–passive; see Figure 20.1).

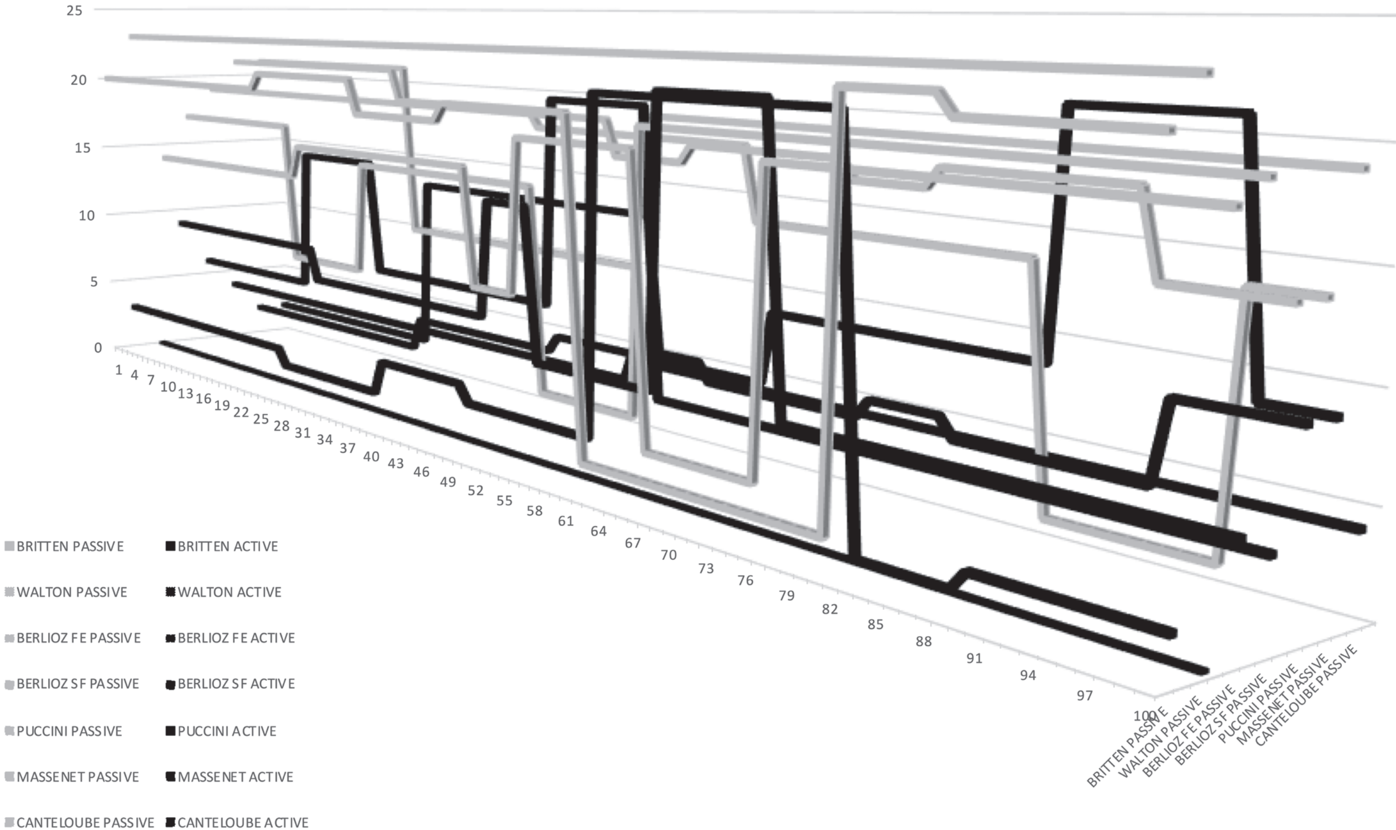


Figure 20.1 A distribution of passive and active imagery types in all pieces of the "Nurturing" programme.

Results show that the following pieces displayed the imagery in a three-part narrative form (passive–active–passive): Britten, *Sentimental Sarabande*; Berlioz, *Flight to Egypt*; Berlioz, *Shepherd's Farewell*; Massenet, *Under the Linden Trees*; and Canteloube, *Berceuse*. The following only displayed passive imagery: Walton, *Touch Her Soft Lips*; and Puccini, *Humming Chorus* (see Figure 20.1).

Discussion and Conclusion

The aim of this study was to evaluate whether music-evoked visual imagery from the compositions in the "Nurturing" programme appeared in a three-part narrative form. Imagery from Britten, *Sentimental Sarabande*; Berlioz, *Shepard's Farewell*; Berlioz, *Flight to Egypt*; Massenet, *Under the Linden Trees*; and Canteloube, *Berceuse*, assumed a three-part narrative form. Imagery in Walton and Puccini did not assume a three-part narrative form.

There are two possible explanations of why the three-part narrative form occurred in the aforementioned pieces. The first is that the occurrence of active and passive imagery is governed by the specific music features in each of the pieces of the programme, which induce the feelings of relaxation or tension and are distributed in each piece in a way that follows the passive–active–passive form. Since Bonny did not choose the pieces of the "Nurturing" programme according to the three-part distribution of their relaxations and tensions, it is likely that the similar three-part blueprint exists in other music pieces too. The second explanation is related to what Fisher (1985) has suggested that the human mind interprets any form of meaningful content in a narrative form. Since music presents a meaningful form of expression, it is possible that certain music pieces are interpreted in a narrative manner.

However, this last argument is invalid if we take a look at the remaining pieces of the "Nurturing" programme: Walton and Puccini's pieces. They did not exhibit the three-part narrative form and hence we cannot conclude that all meaningful content in the form of music is interpreted in a narrative form. The two pieces, however, have two things in common; short length of the pieces (Walton, 1 minute 51 seconds; Puccini, 2 minutes 47 seconds) and the lack of contrast between the successive music features. The short length of the pieces might have prevented the active imagery from developing, but for this there is not enough evidence in the literature, so a definite conclusion cannot be drawn. However, the lack of contrast between the features within the two pieces such as the constant degree of dynamics, and repetitive rhythm, and melodic line might have contributed to the stagnation of imagery (Bruscia et al., 2005). This argument is supported by Lerdahl and Jackendoff's (1983) theory which argues that the tension-relaxation exchange that is conveyed by the music features is the basic force of temporal progression of music. They define the *tension* and *relaxation* in pieces as follows: When a listener hears two music events one after another and experiences the second event as less stable than the first one, the succession will be experienced as *tension*. However, if the second event is more stable than the first, it will be experienced as *relaxation*. Moreover, the perception of tension and relaxation in listeners will depend on their perception of contrast between the two events; the more significant the contrast, the more the alteration will be noticeable to the listeners. Referring back to the results of this study, we can see that the three-part narrative form of imagery only occurred in the pieces that have a pronounced contrast between the consecutive music features, suggesting that the subjective perception of musical relaxations and tensions affects the types of imagery appearing.

Nattiez (2013) also acknowledges the fluctuating pattern of tensions and relaxations present in music using the term *proto-narrative* to indicate the fact that music only "imitates" the real narrative and is thus capable of expressing the blueprint of a story. This is further supported by Daniel Stern's theory (1995); he introduced the term *proto-narrative envelope* to describe how infants organize their experience of time. Early experiences of interaction between a mother and a baby have a beginning, a middle, and an end, and a line of dramatic tension (Stern, 2002) based on the difference between waiting for food (tension) and being fed (relaxation). He called those short interactions proto-narrative envelopes and proposed that they represent a basic unit for constructing the representational world. The results thus indicate that contrasting music features could create the line of dramatic tension in the imagery, similar to the one present in literary narratives, that draws its origins in our earliest experiences of organizing time through "a plot visible only through the perceptual and affective strategies to which it gives rise" (Stern, 2002, p. 180).

In further studies, it would be interesting to examine which music features contributed to the emergence of active and passive imagery. This would provide an opportunity to predict the occurrence of different imagery types and to more accurately and objectively choose an appropriate GIM programme for a client. Therapist's choice of programme would thus depend on the objective qualities of music pieces instead of their intuitive judgement, rendering the process more transparent and measurable. This would further enable the development and acknowledgement of GIM method in a wider psychotherapeutic society. Finally, this study is relevant for both the fields of therapy and psychology, as it paves the way towards two important areas of research: the relationship between the listener's emotion and visual imagery and the origin of our emotional relationship with music (Taruffi & Küssner, this volume).

In conclusion, we can see that music tends to elicit a three-part narrative form of imagery within the framework of GIM. The exchange of passive and active imagery is most likely triggered by the emotional perception of tensions and relaxations created by different music features that need to be examined in more detail.

References

Bonny, H. L. (1989). Sound as symbol: Guided imagery and music in clinical practice. *Music Therapy Perspectives*, *6*(1), 7–10. https://doi.org/10.1093/mtp/6.1.7

Bonny, H. L. (2010). Music psychotherapy: Guided imagery and music. *Voices: A World Forum for Music Therapy*, *10*(3). https://doi.org/10.15845/voices.v10i3.568

Bruscia, K. E., Abbott, E., Cadesky, N., Condron, D., Hunt, A. M., Miller, D., & Thomae, L. (2005). A collaborative heuristic analysis of Imagery-M: A classical music program used in the Bonny Method of Guided Imagery and Music (BMGIM). *Qualitative Inquiries in Music Therapy*, *2*, 1–35.

Bruscia, K. E., & Grocke, D. E. (2002). *Guided imagery and music: The Bonny method and beyond*. Barcelona: Barcelona Publishers.

Campbell, J. (1949). *The hero with a thousand faces* (3rd ed.). Novato, CA: New World Library.

Di Bona, E. (2017). Listening to the space of music. *Rivista di estetica*, (66), 93–105. https://doi.org/10.4000/estetica.3112

Dukić, H., Parncutt, R., & Bunt, L. (2021). Narrative archetypes in the imagery of clients in guided imagery and music therapy sessions. *Psychology of Music*, *49*(2), 287–303. https://doi.org/10.1177/0305735619854122

Field, S. (2007). *The screenwriter's workbook: Exercises and step-by-step instructions for creating a successful screenplay*. New York: A Delta Book.

Fisher, W. R. (1985). The narrative paradigm: An elaboration. *Communication Monographs*, *52*(4), 347–367. https://doi.org/10.1080/03637758509376117

Freytag, G. (1900). *An exposition of dramatic composition and art*. An Authorized Translation From the Sixth German Edition by Elias J. MacEwan, M.A. (3rd ed.), Chicago: Scott, Foresman and Company.

Grocke, D. E. (1999). *A phenomenological study of pivotal moments* [Unpublished Doctoral Dissertation]. The University of Melbourne.

Herman, D. (2009). *Basic elements of narrative*. Hoboken: John Wiley & Sons.

Herman, D., Phelan, J., Rabinowitz, P. J., Richardson, B., & Warhol, R. (2012). *Narrative theory: Core concepts and critical debates*. Columbus: The Ohio State University Press.

Jacobson, E. (1938). *Progressive relaxation*. Chicago: University of Chicago Press.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.

Lewis, K. (1999). The Bonny method of guided imagery and music: Matrix for transpersonal experience. *Journal of the Association for Music and Imagery*, *6*, 63–85.

Margulis, E. H. (2017). An exploratory study of narrative experiences of music. *Music Perception: An Interdisciplinary Journal*, *35*(2), 235–248. https://doi.org/10.1525/mp.2017.35.2.235

Maus, F. E. (1988). Music as drama. *Music Theory Spectrum*, *10*, 56–73. https://doi.org/10.2307/745792

McAdams, D. P. (2006). The problem of narrative coherence. *Journal of Constructivist Psychology*, *19*(2), 109–125. https://doi.org/10.1080/10720530500508720

McKinney, C. H. (1990). The effect of music on imagery. *Journal of Music Therapy*, *27*(1), 34–46. https://doi.org/10.1093/jmt/27.1.34

Nattiez, J. J. (2013). The narrativization of music: Music: Narrative or proto-narrative? *Human and Social Studies*, *2*(2), 61–86. http://dx.doi.org/10.2478%2Fhssr-2013-0004

Robinson, J., & Hatten, R. (2012). Emotions in music. *Music Theory Spectrum*, *34*(2), 71–106. https://doi.org/10.1525/mts.2012.34.2.71

Stern, D. N. (1995). *The motherhood constellation: A unified view of parent-infant psychotherapy*. New York: Basic Books.

Stern, D. N. (2002). *The first relationship: Infant and mother*. Cambridge, MA: Harvard University Press.

Traupman, J. C. (1966). *The New College Latin and English dictionary*. Toronto: Bantam.

VERBI Software. (2015). *MAXQDA 12, software for qualitative data analysis, 1989–2018* [Computer Software]. Berlin: Sozialforschung GmbH.

Wilber, K. (1993). *The spectrum of consciousness*. Adyar: Quest Books.

21 The Image Behind the Sound

Visual Imagery in Music Performance

Graziana Presicce

An increasing number of studies recognize the beneficial effects that mental imagery can bring to the music practice room and beyond—from enhancing performers' musical expressivity (Clark et al., 2012) to regulating performance anxiety (Connolly & Williamon, 2004; Finch et al., 2021; Finch & Oakman, this volume). While these works advance our knowledge of imagery experiences and highlight their potential benefits for music performers, it remains a complex phenomenon to pin down: even when focusing on the visual modality ("seeing" in the absence of a sensory stimulus; Taruffi & Küssner, 2019), we encounter several typologies of visual mental images (Taruffi & Küssner, this volume). Drawing on Christoff et al.'s (2016) research on spontaneous cognition, Taruffi and Küssner (2019) provide a more nuanced understanding of music listeners' experiences of visual imagery through the lens of a new conceptual framework. The aim of this chapter is to adapt specific aspects of Taruffi and Küssner's (2019) framework to describe different typologies of (but not limited to) visual imagery in music performance, drawing on both literature and personal accounts from piano practice sessions. In the first part of this chapter, the author reviews a selection of existing literature on visual imagery and music performance in relation to Taruffi and Küssner's framework, developing a typology of performers' experiences or uses of visual imagery that can be divided into three broad categories: *spontaneous, heuristic*, and *strategic*. In the second part of the chapter, these categories are discussed in relation to qualitative data extracts comprising reflections and accounts of visual imagery derived from the author's personal experiences as a solo pianist and from a brief piano-duet case study.

Typologies of (Visual) Imagery

Drawing on empirical, theoretical, and neuroscientific literature, Christoff and colleagues (2016) developed a conceptual framework that elucidates different types of spontaneous thought. According to this framework, the contents of thoughts and the transition between different mental states are restricted by two general mechanisms: deliberate and automatic constraints. Deliberate constraints are exerted through cognitive control (e.g., deliberately maintaining high levels of concentration while reading a complex book), whereas automatic constraints

DOI: 10.4324/9780429330070-26

operate outside of cognitive control: for instance, sensory or affective salience holding one's attention to a restricted set of information (e.g., despite our efforts to read a book, we are unable to disengage our attention from the sound of a dripping tap or "from a preoccupying emotional concern"; Christoff et al., 2016, p. 719). The framework is dynamic and both types of constraints can vary from weak to strong. Furthermore, when thoughts are freed from both automatic and deliberate constraints, mind-wandering can be experienced. Taruffi and Küssner (2019) later proposed the application of Christoff et al.'s (2016) framework to music listeners' experiences of visual imagery: the way in which music-evoked visual imagery can be influenced by varying degrees of cognitively controlled, deliberate constraints (e.g., the goal-directed imagery experienced during Guided Imagery and Music therapy—see Dukić, this volume; Goldberg, 1995) and automatic constraints (e.g., the affective tone conveyed by the music) highlights "a more nuanced and precise understanding of the several typologies of visual mental images" (Taruffi & Küssner, 2019, p. 65). This new framework also reconciles the diverse range of terms related to visual imagery, marking a clearer distinction between them. For instance, the imagery that emerges from mind-wandering experiences (task unrelated and stimulus independent thoughts, drifting away from an ongoing task; Christoff et al., 2016; Konishi, this volume) and the imagery that emerges from musical daydreams (medium deliberate constraints, but stronger than mind-wandering), defined by Herbert (2018, this volume) as musical experiences marked by a fluctuating attentional focus. Visual mental images are only one of a multitude of components of musical daydreams that are intrinsically multimodal and exhibit simultaneously distributed internal and external attentional focus (Herbert, 2018; Taruffi & Küssner, 2019).

Studies investigating music and visual imagery at times focus on highly specific aspects of the imagery experience or visual association. For instance, visualizing sound dimensions (loudness/pitch/timbre) in terms of colour (hue/saturation/light intensity; Giannakis & Smith, 2001); visually imagining the music's formal structure (Kvifte, 2001); or imagining music in terms of the instrument and its physical properties of action, such as "keyboard imagining" (Baker, 2001). Conversely, other studies exploring performers' experiences consider imagery in the broad, multisensory manner of "seeing," "hearing," and/or "feeling" (see Finch et al., 2021; Gregg et al., 2008), as in various studies exploring mental practice (Holmes, 2005; Ross, 1985). The focus of this chapter rests on *music-related visual imagery* as defined by Taruffi and Küssner (2019, p. 63):

> The mechanism whereby music stimulates internal images in the listener [or performer] consisting of pictorial representations (e.g., natural landscape, colours), embodied image-schemata (e.g., picturing a melodic movement as an ascending or descending image), or complex visual narratives (e.g., similar to that of a movie).

The inclusion of a broader range of examples of visual imagery facilitates a more diverse exploration of the ways in which performers employ or experience different types of visual imagery. It should be noted, however, that while the ensuing discussions focus on visual imagery, this does not exclude the possibility of a concurrent use of other imagery modalities (Floridou, this volume). Imagery is, after all, an "impure phenomenon" (Godøy & Jørgensen, 2001, p. 21; Godøy, this volume); indeed, multimodal imagery "may even be the rule rather than the exception" (Taruffi & Küssner, 2019, p. 62).

Music Performers and Visual Imagery: Theories and Applications

The potential benefits of imagery for music performers have been well-acknowledged across various studies, and even regarded as "increasingly obvious" (Clark et al., 2012, p. 360). Existing research elucidates performers' use of imagery to enhance aspects of their playing or assist the performance preparation process (for comprehensive reviews, see Clark et al., 2012; Connolly & Williamon, 2004)—Table 21.1 provides an overview of selected studies. While these works make reference to visual imagery, this is at times embedded as part of a broader, multisensory imagery discussion, such as "mental practice." Moving back to Taruffi and Küssner's (2019) framework, considering the variety of visual imagery experienced by music performers in relation to automatic and deliberate constraints can also help to clarify different typologies of visual imagery, along with their uses, within the context of music performance. Taruffi and Küssner's framework

Table 21.1 Summary of selected literature exploring imagery in relation to music performance.

Performance Aspects	*Selected Studies*
Ensemble Work	—Action planning and execution (Keller, 2012)
Expressive Playing	—Enhance musical expressivity (Clark et al., 2012) —Facilitating stylistic approaches (Trusheim, 1991) —External focus of attention (Wulf & Mornell, 2008),
Memorization	—Mental representations of the piece/structure (Kvifte, 2001; Mountain, 2001; Saintilan, 2014; Chaffin & Imreh, 2002) —Strategic uses of visual imagery (Ginsborg, 2004)
Mental Rehearsal/ Practice	—Imaginary performance setting (Trusheim, 1991) —Promoting motor anticipation (Bernardi et al., 2013) —Mental practice (Ross, 1985)
Motivational	—Regulating arousal (Finch et al., 2021) —Gaining more interest in the music (Clark et al., 2012)
Performance Anxiety	—Psyching up (Finch et al., 2021) —Mental toughness (Gregg et al., 2008) —Inducing positive states (Connolly & Williamon, 2004)
Tone Production	—Visualizing the ideal sound (Trusheim, 1991) —Adjusting sounds to descriptive words (Auhagen & Schoner, 2001)

indicates *mind-wandering* (Konishi, this volume) as the visual imagery experienced through spontaneous cognition, and *Guided Imagery and Music* (Dukić, this volume) as the visual imagery experienced through goal-directed cognition. In this respect, the present author proposes that performers' visual imagery can be conceptualized into three broad categories: *spontaneous*, *heuristic*, and *strategic* (Figure 21.1 provides a visualization of this extended framework). Each of the three categories is now discussed in turn.

The *spontaneous* category captures the unexpected visual imagery that may emerge during a performance activity. As in music listeners, these experiences include arbitrary or musically unrelated mental images—such as the imagery that emerges from mind-wandering experiences, as pointed out in Taruffi and Küssner's (2019) framework. Arguably, however, there is a need to consider types of imagery that also emerge spontaneously but cannot be identified as mind-wandering, since the latter would imply attention drifting away from the performed task (Smallwood & Schooler, 2015). For instance, the unexpected visual imagery that may emerge from the *magical moments* in Strong Experiences in Music (SEMs: Gabrielsson, 2011): these typically involve a state of *flow* where "music completely dominates one's attention and shuts out everything else" (p. 67). Although performers experience SEMs in a manner similar to music listeners, what is specific to performers is the actual playing, along with their contact with the audience and fellow players (Gabrielsson, 2011, p. 248). Among the accounts reported, a singer describes how "very suddenly," from the first note, he was "inside the song . . . as if the ceiling in the practice room disintegrated, and I was standing there under the stars in the moonlight" (Gabrielsson, 2011, p. 244). The positive relationship between experiences of visual imagery and engagement with the music (being "compelled" and "drawn in") is also supported in two recent empirical studies on music listeners (Presicce, 2019; Presicce & Bailes, 2019)—although the voluntary and spontaneous experiences of imagery were explored conjointly. Occasionally, participants were even surprised to experience the imagery: "I was surprised by how much I did [imagine] . . . but actually as I was there, just focusing on the music, I did" (Presicce, 2019, p. 162). It may be possible that these experiences comprise higher degrees of automatic constraints, as a result of affective salience—such as the strong, intense feelings reported in SEMs (Gabrielsson, 2011). A broader "spontaneous" category could therefore account for both imagery examples of weak deliberate constraints, whether these emerge from mind-wandering experiences or music-driven visual imagery.

The *heuristic* category comprises the mental images that can influence one or more aspects of one's playing, yet the means by which this is achieved is heuristic in nature. Leech-Wilkinson and Prior (2014) define heuristics as "short-cuts based on experience that solve problems too complex to resolve quickly enough using analytical thought" (p. 36). The authors discuss how certain words, while apparently naïve, actually capture rich associations, meanings, and implications acquired through practice (Leech-Wilkinson & Prior, 2014, p. 54). Performers at times skip the detailed decision-making processes through the use of metaphorical ideas (Black, this volume) such as "shape" (e.g., shaping a musical phrase),

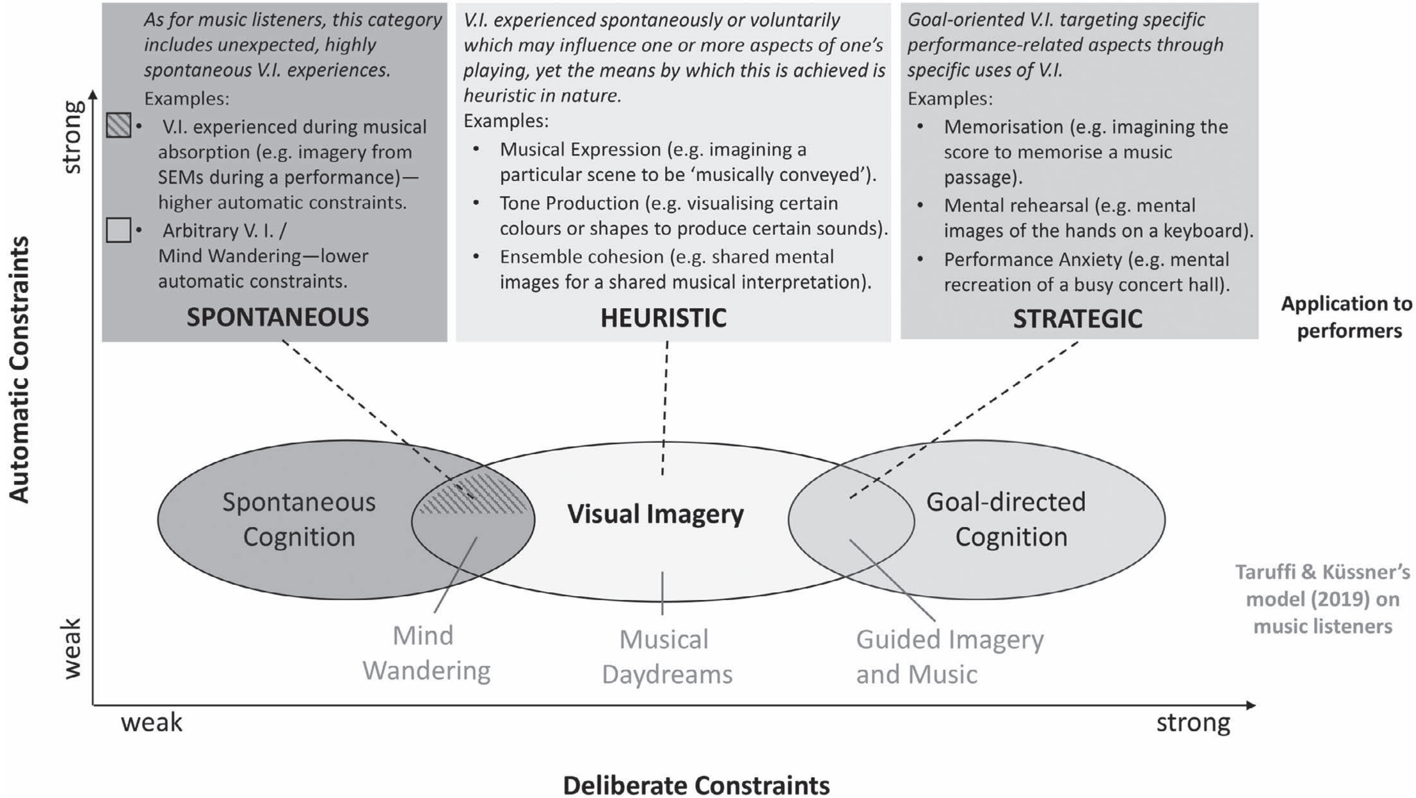


Figure 21.1 Adaptation of selected aspects from Taruffi and Küssner's (2019) framework of music listeners' visual imagery (lower section) to music performers (upper section).

enabling them to create a musically expressive performance (Prior, 2017, p. 228). Although these studies focus on language and metaphors, the same concept could also be applied to the (mental) visual domain. For instance, one brass player reported visualizing his ideal sound in terms of colours: "that shiny bronze or shiny gold—not too gold—not a dull gold" (Trusheim, 1991, p. 142). While ostensibly simple, this process can be highly effective: a mental image can shape one's musical behaviour, alongside the subconscious technical processes that shape that musical behaviour (Kvifte, 2001, p. 222). Thus, the player in the previous example achieved the desired tone through imagery without being consciously aware of the detailed technical adjustments that were required. Heuristic visual imagery lies in the central range of deliberate constraints (Figure 21.1): this is particularly useful, since certain types of imagery may be difficult to dissect in terms of the degree to which they are spontaneous or volitional (for instance, musically responding to an unexpected visual mental image, such as a dark, sinister landscape) and may oscillate between the two during the course of the experience—similar to the fluctuating attentional focus described by Herbert's (2018) "musical daydreams." Arguably, heuristic kinds of visual imagery also differ from *strategic* imagery with regard to the higher degree of creativity involved: according to Christoff et al.'s (2016) framework, creative thinking pertains to a dynamic form of spontaneous thoughts that shifts between the spontaneous generation of new ideas and the critical evaluation of these ideas (more goal-directed). Similarly, performers may engage in a trial-and-error process in search of the "right" mental image to support their musical expression.

Finally, *strategic* visual imagery denotes the goal-oriented mental images especially employed to target specific aspects of performance. In relation to music listeners, Taruffi and Küssner (2019) propose as strong deliberate constraints the visual imagery employed by Guided Imagery and Music (GIM)—a therapeutic practice where patients are invited to share their mental images while being guided through a carefully programmed musical sequence (Goldberg, 1995). In lieu, performers may engage in strategic kinds of visual imagery to improve skills such as music memorization (e.g., imagining the score to memorize a specific passage or the formal music structure; Chaffin & Imreh, 2002; Ginsborg, 2004; Kvifte, 2001; Mountain, 2001; Saintilan, 2014) and mental rehearsal (e.g., visualizing the hands on the instrument and the fingerings used; Bernardi et al., 2013; Ross, 1985; Trusheim, 1991); or to develop coping strategies for performance anxiety (e.g., imagining a busy hall before the performance day; Connolly & Williamon, 2004; Finch et al., 2021; Gregg et al., 2008). Hence, strategic visual imagery targets specific steps towards achieving the intended performance goal.

As a result of the fluid, complex nature of imagery, it is important to point out that imagery experiences may indeed overlap or fluctuate between these three categories. Nonetheless, the adapted framework constitutes an initial attempt towards contextualizing performers' visual imagery in relation to recent research, and in line with Taruffi and Küssner's (2019) work, with a view to developing a more nuanced understanding of various types of music-evoked visual imagery. The next part of the chapter provides examples of visual imagery in a performance preparation context. These are drawn from a brief piano-duet case study where the author was one of the pianists, as well as extracts from the author's visual imagery in selected piano solo works.

Pianists and Visual Imagery—Examples

Participants

The extracts of qualitative data presented later emerge from a broader study (Presicce, 2019) exploring visual imagery in music listeners. Qualitative accounts from the piano duet involved the author (Pianist I, female, age = 26); and a second professional pianist (Pianist II, female, age = 34).

Procedure

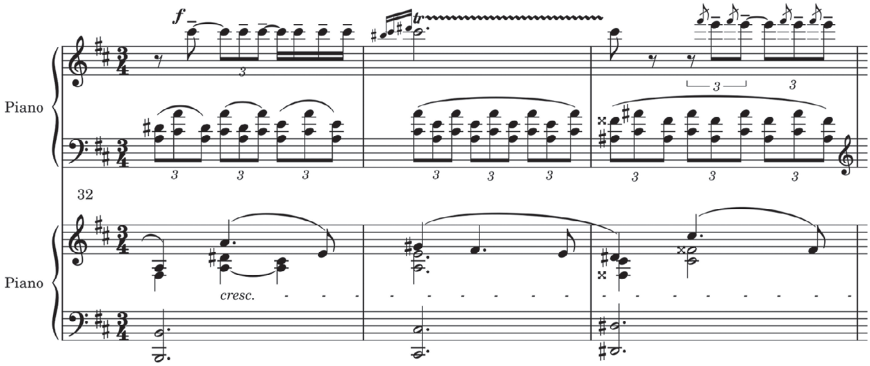
The two pianists attended together one rehearsal (2 hours) in preparation for a recording of Rachmaninov's *Fantaisie-Tableaux*, Op. 5 (movements 2–4; Table 21.2); they had never previously performed together. After the joint rehearsal, the following instructions were provided: "Please use the following form to keep a record of any visual imagery that spontaneously occurs or is intentionally created during your private practice of this ensemble's works." The instructions also asked them to point out, whenever possible, at what point in the music the imagery emerged. The annotations were collected one month later, prior to the recording session. Both pianists produced one imagery report per piece; details of when the imagery occurred over time, however, were not provided. The three movements rehearsed present particularly suggestive titles and their accompanying epigraphs cite extracts of poems by Byron, Tioutchef, and Khomyakof. Pianist II reported her imagery more anecdotally, whereas Pianist I presented her imagery in more detailed prose. Since both pianists reported imagery pertinent to the musical material, the following examples focus on expressivity, tone production (*heuristic imagery*), and memorization techniques (*strategic imagery*).

Heuristic Visual Imagery

In Rachmaninov's *La Nuit . . . L'Amour* (second movement), both pianists reported love- and bird-related visual imagery, themes that also appear in the

Table 21.2 Details of the music pieces used as auditory stimuli for the study, selected from Rachmaninov's *Fantaisie-Tableaux*, Op. 5, for two pianos.

Composer	*Piece*	*Date of Composition*	*Track Length*
S. Rachmaninov	II. *La Nuit . . . L'Amour*	1893	6′23″
	III. *Les Larmes*		6′13″
	IV. *Pâques*		2′57″



No.2 *La Nuit... L'Amour*
It is the hour when from the boughs
The nightingale's high note is heard;
It is the hour when lovers' vows
Seem sweet in every whisper'd word;
And gentle winds, and waters near,
Make music to the lonely ear.
Byron

Figure 21.2 Bars 32–34 from *La Nuit . . . L'Amour,* Rachmaninov.
Source: The work's accompanying epigraph (translation by Boosey & Hawkes)

work's epigraph. This is in line with previous research suggesting that contextual information about a musical work promotes visual imagery related to those descriptions (Vuoskoski & Eerola, 2013) and, as explored later, the way such information can bear a direct influence on the musical shaping used by performers (Prior, 2017). A vivid instance of Pianist I's bird imagery emerged with the appearance of trills, single-note repetitions and acciaccaturas (bars 32–4, Figure 21.2). Pianist I described how she adjusted her playing as a result of the mental image:

> When I have these figures, I try to shape my playing to how a bird's voice would sound; to achieve this, it feels like the music at this point does not require to be rhythmically overly strict (obviously still keeping the overall rhythm, but notes do not have to be metronomic). I try to make these singing calls bright and emerging from the texture.

Bird imagery therefore encouraged Pianist I's rhythmic flow to deviate from a strict pulse and to adjust the dynamic balance between the two hands. The emergence of the right hand also appears as an attempt to allow that particular musical material to detach from the overall textures: Pianist I's triplets (left-hand) are required to maintain a steady beat to enable a sensible synchronization with Pianist II which, nonetheless, has the main musical theme. It should be noted that although the earlier description highlights the sound adjustments undertaken by the pianist, these were nonetheless prompted by a visual mental image. While a broader connection was made from the trills to bird calls, the sound association—at least on a conscious level—emerged from knowledge, rather than auditory imagery (knowing how birds sound and linking these to birds, rather than actively imagining the sound of bird calls while playing). Hence, trills initially prompted visual imagery of birds, and the playing was subsequently adjusted to convey the aural qualities of bird calls. Once again, this highlights the complexities involved in deciphering and expressing imagery experiences. A further example is described by Pianist I in relation to *Pâques* (fourth movement, see Supplementary Material):

> [*Pâques*] includes the idea of joy and . . . a kind of light so bright that its colour is intense and luminous. When the theme returns, doubled by further octaves from the end of bar 8, these conjure up larger bells: they have richer sonorities; hence, in my mind, the image now seems on a larger scale. For this, the attack to the keys is from further a distance, striking from above but still aiming for a sonorous sound than a harsh one (at this end, the hand strikes flat on the keys' surface with a fast attack, yet without overly "digging into the keys").

While the visual imagery of the two performers frequently shared common themes across the pieces, at times these differed. For instance, in bar 98 of *La Nuit . . . L'Amour*, Pianist I describes her visual imagery as somewhat difficult to verbally depict but centred around a theme of "love" and "warmth." She describes "reaching/getting closer to the desired person/half" but without "a clear idea whether any of the [piano] parts 'characterized' a male or female . . . but there is nonetheless attraction between the two parts." As a result, Pianist I attempted "to create warmth rather than a 'menacing' tone, due to its low register," by tailing off towards the end of the phrase. She adds: "it's a passage that can easily be played sending the wrong message, due to its very low register." The author's personal technical translation of the aforementioned imagery results in a weighty approach to the keys, yet with a soft attack: "digging" in the keyboard with the weight of the arm, maintaining close contact with the keys before pressure is applied, and using a larger surface area of the fingers. The slight diminuendo at the end of the phrase is particularly important here to avoid the menacing effect described. However, this imagery contrasts with that experienced by Pianist II, which diverges from the work's epigraph. When the same theme appears in her part, Pianist II describes a "roaring lion, parading his territory." Her shaping of the phrase at this point also differs, applying a small *crescendo* and *accelerando* towards the end of the phrase and using a slightly broader dynamic range. While subtle, such differences can convey different interpretative approaches to this passage. In this short study, the rehearsal time available to develop as an ensemble was particularly restricted. Yet there is an under-explored potential for visual imagery within the ensemble context. A study by Auhagen and Schoner (2001) demonstrates that musicians not only adapt their sound quality to match given verbal attributes (e.g., bright or dark), but they also show consistency in the variation of their playing parameters. Visual imagery may trigger musical adjustments in a similar manner. Therefore, reconciling differing visual mental images could help in working towards shared interpretational goals, potentially improving ensemble cohesion, and giving rise to creative teamwork towards achieving performance excellence.

Strategic Visual Imagery

The visual imagery discussed thus far links, in one way or another, with the performers' interpretations of the musical work. However, performers may also strategically employ visual imagery for specific goals. For instance, building on the potential of visual imagery within the ensemble context, a study by Keller (2012) indicates how the deliberate utilization of visual, motor, and auditory imagery can assist the planning and execution of performers' actions, providing better control over parameters such as timing and articulation. The deliberate use of visual imagery to assist music memorization has also been widely acknowledged (Ginsborg, 2004; Kvifte, 2001; among others). Reflecting on the author's personal experiences of performing memorized music, visual imagery enables a kind of "mental mapping" of the musical work, providing the visual support that would normally be supplied by the music score. These kinds of mental images tend to be conjured up only briefly and at specific moments in the piece, offering "a convenient way to navigate from one section to another" (Mountain, 2001, p. 277). The use of such "referential points" aligns with Ginsborg's (2004) recommendations, which warn against the dangers of relying solely on visual memory: deliberately planned visual memory, instead, serves more effective as "reliable cues" during the course of a performance (p. 130).

An example of this emerges from the author's visual imagery in Poulenc's *Caprice Italien* (*Napoli suite*). The entrance of the tarantella section (bar 31) presents a considerable leap from the lower to the higher register of the keyboard. The fast pace of the music requires an especially deft movement of the hands to approach the initial notes of bar 31 within a particularly short time-frame, and to position the fingers ready for the notes which follow. While approaching this passage, it is often the image of the hands' positions on those keys that is conjured up during playing (Figure 21.3). This mental image is only brief, yet long enough to ensure the right keys are targeted. As in Chaffin and Imreh (2002), performance cues are utilized as retrieval cues: the image of the hand position here is particularly ingrained as a personal representation of this section, alongside the pictorial kinds of imagery linked to its tarantella element. As pointed out by Saintilan (2014, p. 309), such mental maps—unique to the individual—bring to light an "inner representation of sound, activity, and images used by the musicians to structure their playing."

Figure 21.3 Position of the hands at bar 31 of Poulenc's *Caprice Italien*, third movement of *Napoli* suite.

Closing Remarks

Based on Christoff et al.'s (2016) research on spontaneous thoughts, the present chapter applied Taruffi and Küssner's (2019) framework of music listeners' visual imagery experiences to music performers, supported by selected examples of qualitative data derived from a brief piano-duet case study and the author's personal imagery experiences in relation to piano solo works. The adaptation of the framework suggests that performers' experiences of visual imagery may be gathered into three broad categories: *spontaneous*, *heuristic*, and *strategic*, respectively, ranging from weak to strong deliberate constraints. Notably, the examples of visual imagery presented referred to both *heuristic* and *strategic* imagery; neither pianist reported *spontaneous* imagery experiences such as mind-wandering. It may be the case that mental images not related to the music were deemed irrelevant to the task and, therefore, not reported. While the idiosyncratic nature of these experiences is acknowledged, these examples nonetheless provide a glimpse into the ways in which visual imagery can go beyond a "mere visualization," bearing on technical and expressive aspects of one's playing—a key distinction from Taruffi and Küssner's (2019) framework for music listeners. The aforementioned examples also align with existing literature. For instance, visual imagery can be employed strategically and selectively to be most effective in music memorization (Ginsborg, 2004; Kvifte, 2001), whereas creative imagery experiences can act as heuristics that influence the playing, in a manner similar to verbal metaphors (Auhagen & Schoner, 2001; Prior, 2017). Overall, visual imagery can be a powerful and versatile tool available for performers. Nonetheless, abundant scope remains for further investigation.

References

Auhagen, W., & Schoner, V. (2001). Control of timbre by musicians: A preliminary report. In R. I. Godøy & H. Jørgensen (Eds.), *Musical imagery* (pp. 202–217). Milton Park: Routledge. https://doi.org/10.4324/9780203059494

Baker, J. (2001). The keyboard as basis for imagery of pitch relations. In R. I. Godøy & H. Jørgensen (Eds.), *Musical imagery* (pp. 251–270). Milton Park: Routledge. https://doi.org/10.4324/9780203059494

Bernardi, N. F., De Buglio, M., Trimarchi, P. D., Chielli, A., & Bricolo, E. (2013). Mental practice promotes motor anticipation: Evidence from skilled music performance. *Frontiers in Human Neuroscience*, *7*, 451. https://doi.org/10.3389/fnhum.2013.00451

Chaffin, R., & Imreh, G. (2002). Practicing perfection: Piano performance as expert memory. *Psychological Science*, *13*(4), 342–349. https://doi.org/10.1111/j.0956-7976.2002.00462.x

Christoff, K., Irving, Z. C., Fox, K. C., Spreng, R. N., & Andrews-Hanna, J. R. (2016). Mind-wandering as spontaneous thought: A dynamic framework. *Nature Reviews Neuroscience*, *17*(11), 718–731. https://doi.org/10.1038/nrn.2016.113

Clark, T., Williamon, A., & Aksentijevic, A. (2012). Musical imagery and imagination: The function, measurement and application of imagery skills for performance. In D. J. Hargreaves, D. E. Miell, & R. A. R. MacDonald (Eds.), *Musical imaginations: Multidisciplinary perspectives on creativity, performance and perception* (pp. 351–368). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199568086.003.0022

Connolly, C., & Williamon, A. (2004). Mental skills training. In A. Williamon (Ed.), *Musical excellence: Strategies and techniques to enhance performance* (pp. 221–245). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198525356.003.0012

Finch, K. K., Oakman, J. M., Milovanov, A., Keleher, B., & Capobianco, K. (2021). Mental imagery and musical performance: Development of the Musician's Arousal Regulation Imagery Scale. *Psychology of Music*, *49*(2), 227–245. https://doi.org/10.1177/0305735619849628

Gabrielsson, A. (2011). *Strong experiences with music*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199695225.001.0001

Giannakis, K., & Smith, M. (2001). Imaging soundscapes: Identifying cognitive associations between auditory and visual dimensions. In R. I. Godøy & H. Jørgensen (Eds.), *Musical imagery* (pp. 161–180). Milton Park: Routledge. https://doi.org/10.4324/9780203059494

Ginsborg, J. (2004). Strategies for memorizing music. In A. Williamon (Ed.), *Musical excellence: Strategies and techniques to enhance performance* (pp. 123–142). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198525356.003.0007

Godøy, R. I., & Jørgensen, H. (2001). Overview. In R. I. Godøy & H. Jørgensen (Eds.), *Musical imagery* (pp. 1–4). Milton Park: Routledge. https://doi.org/10.4324/9780203059494

Goldberg, F. S. (1995). The Bonny method of guided imagery and music. In T. Wigram, B. Saperston, & R. West (Eds.), *The art and science of music therapy: A handbook* (pp. 112–128). Milton Park: Routledge.

Gregg, M. J., Clark, T. W., & Hall, C. R. (2008). Seeing the sound: An exploration of the use of mental imagery by classical musicians. *Musicae Scientiae*, *12*(2), 231–247. https://doi.org/10.1177/102986490801200203

Herbert, R. (2018, May 16–19). *Everyday musical daydreams and kinds of consciousness* [Paper presentation]. KOSMOS Workshop "Mind Wandering and Visual Mental Imagery in Music", Berlin, Germany.

Holmes, P. (2005). Imagination in practice: A study of the integrated roles of interpretation, imagery and technique in the learning and memorisation processes of two experienced solo performers. *British Journal of Music Education*, *22*(3), 217–235. https://doi.org/10.1017/S0265051705006613

Keller, P. E. (2012). Mental imagery in music performance: Underlying mechanisms and potential benefits. *Annals of the New York Academy of Sciences*, *1252*(1), 206–213. https://doi.org/10.1111/j.1749-6632.2011.06439.x

Kvifte, T. (2001). Images of form: An example from Norwegian Hardingfiddle Music. In R. I. Godøy & H. Jørgensen (Eds.), *Musical imagery* (pp. 219–236). Milton Park: Routledge. https://doi.org/10.4324/9780203059494

Leech-Wilkinson, D., & Prior, H. M. (2014). Heuristics for expressive performance. In D. Fabian, R. Timmers, & E. Schubert (Eds.), *Expressiveness in music performance: Empirical approaches across styles and cultures* (pp. 34–57). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199659647.003.0003

Mountain, R. (2001). Composer and imagery: Myths and realities. In R. I. Godøy & H. Jørgensen (Eds.), *Musical imagery* (pp. 272–288). Milton Park: Routledge. https://doi.org/10.4324/9780203059494

Presicce, G. (2019). *The role of engagement and visual imagery in music listening* [University of Hull]. Ethos. https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.815107

Presicce, G., & Bailes, F. (2019). Engagement and visual imagery in music listening: An exploratory study. *Psychomusicology: Music, Mind, and Brain*, *29*(2–3), 136–155. https://doi.org/10.1037/pmu0000243

Prior, H. (2017). Shape as understood by performing musicians. In H. Leech-Wilkinson & H. Prior (Eds.), *Music and shape* (pp. 216–241). Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780199351411.003.0014

Ross, S. L. (1985). The effectiveness of mental practice in improving the performance of college trombonists. *Journal of Research in Music Education*, *33*(4), 221–230. https://doi.org/10.2307/3345249

Saintilan, N. (2014). The use of imagery during the performance of memorized music. *Psychomusicology: Music, Mind, and Brain*, *24*(4), 309–315. https://doi.org/10.1037/pmu0000080

Smallwood, J., & Schooler, J. W. (2015). The science of mind wandering: Empirically navigating the stream of consciousness. *Annual Review of Psychology*, *66*(1), 487–518. https://doi.org/10.1146/annurev-psych-010814-015331

Taruffi, L., & Küssner, M. B. (2019). A review of music-evoked visual mental imagery: Conceptual issues, relation to emotion, and functional outcome. *Psychomusicology: Music, Mind, and Brain*, *29*(2–3), 62–74. https://doi.org/10.1037/pmu0000226

Trusheim, W. H. (1991). Audiation and mental imagery: Implications for artistic performance. *The Quarterly*, *2*(1–2), 138–147.

Vuoskoski, J. K., & Eerola, T. (2013). Extramusical information contributes to emotions induced by music. *Psychology of Music*, *43*(2), 262–274. https://doi.org/10.1177/0305735613502373

Wulf, G., & Mornell, A. L. (2008). Insights about practice from the perspective of motor learning: A review. *Music Performance Research*, *2*, 1–25.

22 Multimodal Perception in Selected Visually Impaired Pianists

Towards a Conceptual Framework

Anri Herbst and Silvia van Zyl

People's daily interactions with their environments involve a set of complex systems influenced by factors such as the integration of multiple senses, the ability to imagine objects and emotions, and the use of metaphors and conceptual comprehension through bodily experiences (Harrar et al., 2018; Loeffler et al., 2016). When particular human senses are absent or compromised, corresponding cognitive processes are specifically honed to enable interacting with social environments.

Despite large numbers of visually impaired (VI) people worldwide,[1] research on VI musicians is minimal, with little or no reference to mental imagery (cf. Baker & Green, 2017; Ockelford & Matawa, 2009; Wootton, 2005), which forms part of the focus of our study. While this issue has attracted some attention in the case of sighted pianists (Endestad et al., 2020; Loimusalo & Huovinen, 2018), mental imagery in VI pianists received almost none apart from reference to tactile acuity (Legge et al., 2019).

Approaching the study from within a bio-psycho-social context, we view VI pianists as *musicians* who navigate their internal and external environments to function in a performance-oriented sonic world. We investigated a small cohort of VI South African pianists in search of emergent themes with regard to the phenomenon of mental imagery, which we viewed through the lens of relevant cognitive-based literature and applied principles of embodied cognition theory (ECT) and dynamic systems theory (DST). At the end of the chapter, we present a conceptual framework that consolidates the emergent themes from the qualitative findings with relevant cognitive-based literature and supporting theories. The emergent themes from the qualitative study in combination with cognitive-based literature and applied principles of DST and ECT theories culminate in the proposed conceptual framework (cf. Charmaz, 2014, pp. 310–314).

Defining Mental Imagery

Mental imagery as a concept within the context of this study partly developed from musicological contexts, mostly referring to sighted musicians, using terms such as *inner hearing*, *auditory imagery*, *audiation*, *auralization,* and *anticipatory auditory imagery* (Covington, 2005; Gordon, 2012; Karpinski, 2000).

DOI: 10.4324/9780429330070-27

Andrianopoulou (2019) and Herbst (2021) only recently included further facets of embodied cognition such as kinaesthesia (Kim, this volume) as representative of other kinds of imagery not formerly clearly acknowledged.

Broader approaches to mental imagery emerging during music listening (auditory, visual, and/or tactile perception) and reaching beyond music listening and performance-related contexts have attracted increasing scholarly attention (Taruffi & Küssner, 2019).

Mental imagery as a construct substantially informs the formation of our conceptual framework. We define *mental imagery* as mental experiences resulting from multidimensional perceptual processes occurring in the absence of auditory or visual stimuli, assimilating images rooted in the combination of sensory experiences (i.e., through embodiment). These images could originate from past lived experience, as well as forward-and-backward journeying through a sonic landscape. This journey involves elements of listening, tactile imagery, absorption (Vroegh, this volume), mind-wandering (Konishi, this volume), anticipatory imagery, metaphor, triggering of affect, non-musical associations and images (referentialist meaning and episodic memory; Jakubowski, this volume), and theoretical constructs (absolutist meaning, semantic memory). Because of a lack of consensus of the term "mental imagery," we drew on information by a number of scholars to present a multidimensional definition that influences and consolidates various components of the phenomenon (Baddeley, 2012; Cox, 2017; Gabrielsson & Lindström, 2010; Huron, 2006; Johnson & Larson, 2003; Kosslyn et al., 2006; Mast et al., 2012; Meyer, 1956; Moulton & Kosslyn, 2009; Nanay, 2018; Schmidt & Blankenburg, 2019; Vroegh, 2019).

Vroegh's (2019, this volume) research on levels of absorption focused on listening to music and altered states of attention in the listener. Vroegh's study identified the key aspects of *mind-wandering*, *meta-awareness*, *zoning in* and *tuning in* with their binary opposites of *zoning out* and *tuning out*. His levels correspond to some extent with Gordon's six levels of audiation (2012), which also imply a forward and backward mental movement through a musical work. Audiation is not seen as a forward-moving linear process.

Senses

Research with improved brain-mapping technology of visual mental imagery (Pearson et al., 2015), when applied to VI musicians, implies that the ability to imagine objects, scenery, sounds, gustatory and olfactory sensations, alongside phenomena such as synaesthesia,[2] tactile and/or kinaesthetic experiences, while relevant for sighted musicians, is particularly significant to this study as the visual sense is compromised.

Opposing the Aristotelian view of five discrete senses, Durie (2005) proposed 33 senses, of which proprioception, the ability to determine positioning and orientation of extremities such as the hand (Gosselin-Kessiby et al., 2009), and equilibrioception, a sense of balance and posture (Parreira et al., 2017), are essential to this study. Although posture, balance, and keyboard orientation, compared to that of sighted individuals, is less developed in VI, these could be improved through music instruction (Wootton, 2005). VI individuals also employ echolocation, the harnessing of sound reflecting off objects for navigational purposes, in this case showing superior performance compared to sighted individuals (Schenkman & Nilsson, 2010). Cattaneo and Vecchi (2011), Cox (2017) and Shapiro (2019) validated an expansion of the Aristotelian five sense to eight.

Qualitative Study

Four VI Western classical pianists and one sighted piano teacher, who offered an alternative perspective on the learning experiences of VI piano students over a 20-year period (age range 19 to 67 years, M = 42.2), formed part of the qualitative study that made use of purpose-based sampling.[3] This study aimed to investigate the multifaceted phenomenon of mental imagery as experienced by a small number of VI pianists in an attempt to construct a foundational conceptual framework as a portal for further investigation.

Situated in hermeneutic phenomenology (Laverty, 2003), open-ended semi-structured interviews were conducted during June and August 2019 with participants on their levels of impairment and performance; ability to play by ear; thoughts occurring during learning and performing a new piece; experiences of mental imagery and affect during music performances and listening; the appearance of non-musical associations with music theoretical concepts; and mental imagery linked to music notation systems. Based on principles of narrative enquiry, questions were asked on responses elicited when performing and listening to music without disclosing the intention of the study allowing the capturing of the largest possible spectrum of answers.

Participants

The profile of each participant is set out and includes a brief acknowledgement of their type and level of visual impairment as this has direct bearing on the emphasis placed on certain modalities within the evolving framework:

- Participant A: A 19-year-old female, congenitally blind (Leber congenital amaurosis, tunnel vision, light perception, and colour blindness), Grade VIII Piano, Trinity College London (TCL);[4]
- Participant B: A 33-year-old female, partially sighted (congenital bilateral coloboma of the iris, retina, macula and optic nerve, nystagmus), professional pianist and singer, postgraduate music student;
- Participant C: A 34-year-old male, partially sighted until the age of 10, thereafter blind (glaucoma), Grade VI Piano, TCL;
- Participant D: A 67-year-old male, partially sighted (surgical removal of congenital cataracts, amblyopia in right eye, lost left eye in one of the operations, visual acuity 6/60, 10%); PhD graduate, professional musician

(Western classical, jazz piano, and church organ), lecturer at a South African tertiary institution in applied music theory, former head of department;
- Participant E: A 58-year-old female, fully sighted, PhD graduate with topic on teaching VI musicians; has been teaching piano, recorder and general Arts and Culture for 20 years to VI musicians.

Findings

Analysis of audio-visual transcriptions employed grounded theory with initial and axial manual coding (Charmaz, 2014). Six themes reflecting participants' responses emerged from the data analysis:

1 Complexity of braille notation;
2 Focus on music theoretical aspects of a work;
3 Use of metaphor;
4 Use of non-visual modalities;
5 Experiences of affect;
6 Mind-wandering.

Braille

All participants discussed the fragmentary and cumbersome nature of the learning process when using braille notation (cf. Park & Kim, 2014), as information gained is limited to the *tactile* sense, thus hampering access to music material. One hand is used to read short sections with the fingertips, while the other hand plays what is being read on the piano. Participant E focuses on longer phrasing. Sometimes the whole hand brushes over the braille score to form impressions of the grand stave and longer sections.

Even though Participant D mastered braille during his school years, he favoured other approaches such as listening to slowed-down tape recordings. He later adopted Western staff notation as a back-up system, but can still read on only one melodic line at a time. He favours an aural-based learning process grounded in music theoretical knowledge. Participant C conjured up music theoretical aspects captured in braille notation exclusively, but did not mention if this occurred in real time during a performance.

Music Theoretical Aspects of a Work

All participants reported focused attention on music theoretical structures (pitch, harmony, rhythm, timbre, texture, form, etc.) during learning, performing, and listening to a work. Only Participant D referred to using mental imagery present in real-time music theoretical analysis during performing and active listening, implying the consistent presence of the entire formal structure at the back of the mind, thus evoking the sonic landscape presented in Figure 22.1.

Music theoretical structures are important as an emergent theme as understanding and recognizing these structures could conjure up metaphors, allegories, and affects, as explained later in this chapter.

Metaphor

Visual, verbal, space-related, and gesture-related metaphors are often used to link theoretical knowledge and stylistic interpretation of works.

Four participants conjured up visual metaphors to comprehend a work. Some VI pianists associate letters with braille cells, but Participant E was unsure whether her students linked braille cells to specific objects (iconic metaphor).

Programme music evokes pictures (visual metaphors) and stories (allegories). Referring to performing Debussy's *Feuilles morte*, Participant B said: "You would feel as if you are walking in the woods and you got layers of dead wet leaves around you, which you can also smell." She explained that the texture of a fabric could evoke pictures that would assist her in interpreting the mood of a work: "Chiffon would be light, flowy, soft and it's very malleable . . . and could be associated with an ethereal mood," indicating the presence of multisensory metaphors (personal communication, June 28, 2019).

Allegory, a subset of metaphors, played a role in creating stories to interpret a work. Participant E argued that stories should never be imposed, as students' conceptualizations are based on individual sound and tactile experiences. Participants A, B, and E created stories to comprehend 20th- and 21st-century works, which they described as "untidy" as opposed to "tidy" for earlier style periods. Participant E often required learners to write stories incorporating all the senses, later turning them into sound stories with dramatic effects.

Non-Visual Modalities

Participant D associates colour with pitch material, speculating that his sense of absolute pitch may influence his "synaesthesia," which leads him to associate colours with tonalities and days of the week. However, "various types of rhythms do not really elicit specific colourations for me" (personal communication, August 23, 2019).

Three participants referred to "feeling" chords on the piano (spatial orientation, embodied cognition) when asked how they find chords on the instrument. One partially sighted participant had an olfactory association with broccoli when performing or listening to works by Scarlatti, whose music she "didn't really" enjoy.

Participant B referred to echolocation in navigating her physical environment. It is uncertain how this influences musical performance. Participant E indicated that equilibrioception is weak in VI persons, often seen in pianists hooking their legs around the piano chair and slouching, but this improves with instruction over time. It could be argued that all pianists have to develop equilibrioception, muscle memory and a phylogenetic mental image of *correct* posture at the instrument (embodied cognition).

Regarding proprioception, Participant B described difficulty accomplishing large jumps on the piano, indicating poor spatial orientation. Participant D confirmed this by noting that there are more blind organists than pianists, as playing the organ requires fewer jumps.

Affect

Most participants experienced elicited emotions during performance and listening. Participant A explained (paraphrased and translated from her mother tongue Afrikaans into English):

> The Mozart that I am currently learning starts with this dark first octave, . . ., which [angrily stamping foot on floor] makes me angry. And then the music goes softer, and you get to that "nice" little section where you feel contented and light. I am going through different emotions in one piece.
>
> (Personal communication, August 7, 2019)

As church organist, Participant D would sometimes manipulate the material to create a "dignified" way to communicate a specific affect by, for example, playing a Bach Chorale Prelude in a slower tempo during communion. He views sacred music as a way of worship, presenting "in sound what [he is] experiencing in the actual service." This could be a way of conjuring up sublime emotions and images in himself and the congregation.

Mind-Wandering

Participant D described complete immersion in the music during performances: "I am internalising things much more than externalising things. . . . I experience [structure] all the time, because I am always aware of what I am playing" (Personal communication, August 23, 2019).

His listening and playing processes are deeply connected. Real-time analysis always forms part of his active involvement in listening, where he imagines himself as being part of the performance, as if he were playing it himself, thinking about the performance criteria required. However, "sometimes I would flip out of the analysis modes into a creative world that makes me think about something completely different." When very familiar with a work, "you can allow your mind to wander a little bit" as there are "layers of things."

Discussion of Findings Through the Lens of Embodied Cognition Theory and Dynamic Systems Theory

Embodied cognition theory (ECT) (Korsakova-Kreyn, 2018; Matya, 2016; Perlovsky, 2015) explains the cyclical brain processes that shape human perception. The human nervous system consists of the central nervous system (brain and spinal cord) interacting with the peripheral nervous system (PNS—nerves to and from CNS) (Patestas & Gartner, 2016) *as an internal active agent of cognition* (Shapiro, 2019; see Figure 22.1).

ECT holds promise for understanding the linking of bodily experiences with mental images of VI musicians, who need to rely more strongly on alternative senses to compensate for vision loss. Cox (2017) proposes two interlinked motoric-driven processes as components of music cognition: *mimetic motor action* (MMA, actual movement) and *mimetic motor imagery* (MMI, imagination of movement). The motoric and somatosensory nature of the body's interactions with its environment form part of perceptual and cognitive processes underlying all dimensions of musical engagement (Cox, 2017; Geeves & Sutton, 2014).

The brain–body connection is fundamental to ECT, the understanding of which may be augmented to a macro level when complemented by dynamic systems theory (DST), a branch of systems theory (Thelen, 2005) that is widely utilized in multiple fields including transdisciplinary research (Hiver et al., 2021) and functions within all environments of human development and experience. In DST, multicausality is a fundamental feature of complex non-linear living systems, which interact with their environments from a biological cellular level to the highest level of cognition (Perone & Simmering, 2017) as displayed in Figure 22.1 in the box labelled "Perceptual Processing."

Understanding embodied cognition as part of a dynamic system within the ecological environment can help define the creative processes that VI pianists use to navigate their changing physical and mental landscapes, especially but not exclusively through sensorimotor cognition.

The aesthetic approach propounded by Johnson (2007, p. 264) posits that "[p]hilosophy needs a visceral connection to lived experience," which connects with Gallagher's notion that "[a]ffect is deeply embodied" (Gallagher, 2014, p. 16). The use of *metaphor*, defined as "understanding and experiencing one kind of thing in terms of another" (Lakoff & Johnson, 2011, p. 5), could therefore be utilized in expressing imagination and sensorimotor-driven meaning, phenomena also but not exclusively evoked by affect. Visual, verbal, space-related, and gesture-related metaphors are often used to link theoretical knowledge and stylistic interpretation of works: "metaphorization is the basic mechanism in conceptualization of music elements" (Petrović & Golubović, 2018, p. 631).

An understanding of embodied cognition can be extended into a synergistic model where neural, bodily, and environmental fluctuating systems influence one another reciprocally (Van der Schyff et al., 2018). Chemero (2016) refers to Merleau-Ponty's example of blind persons exploring their environments through the tip of a cane, representing a temporary synergistic system that moves beyond the body, understood as extended cognition. The synergy formed between all pianists (VI included) and the piano entails the piano becoming an extension of the player's hands and feet, aiding their navigation of the musical landscape via sensorimotor empathy. Chemero uses the term "Einfühlung" for this, referring to the experience of *feeling into* the tool (the piano), thus "feeling-into 'sensory-motor empathy'" (Chemero, 2016, p. 144). The synergy between pianists and

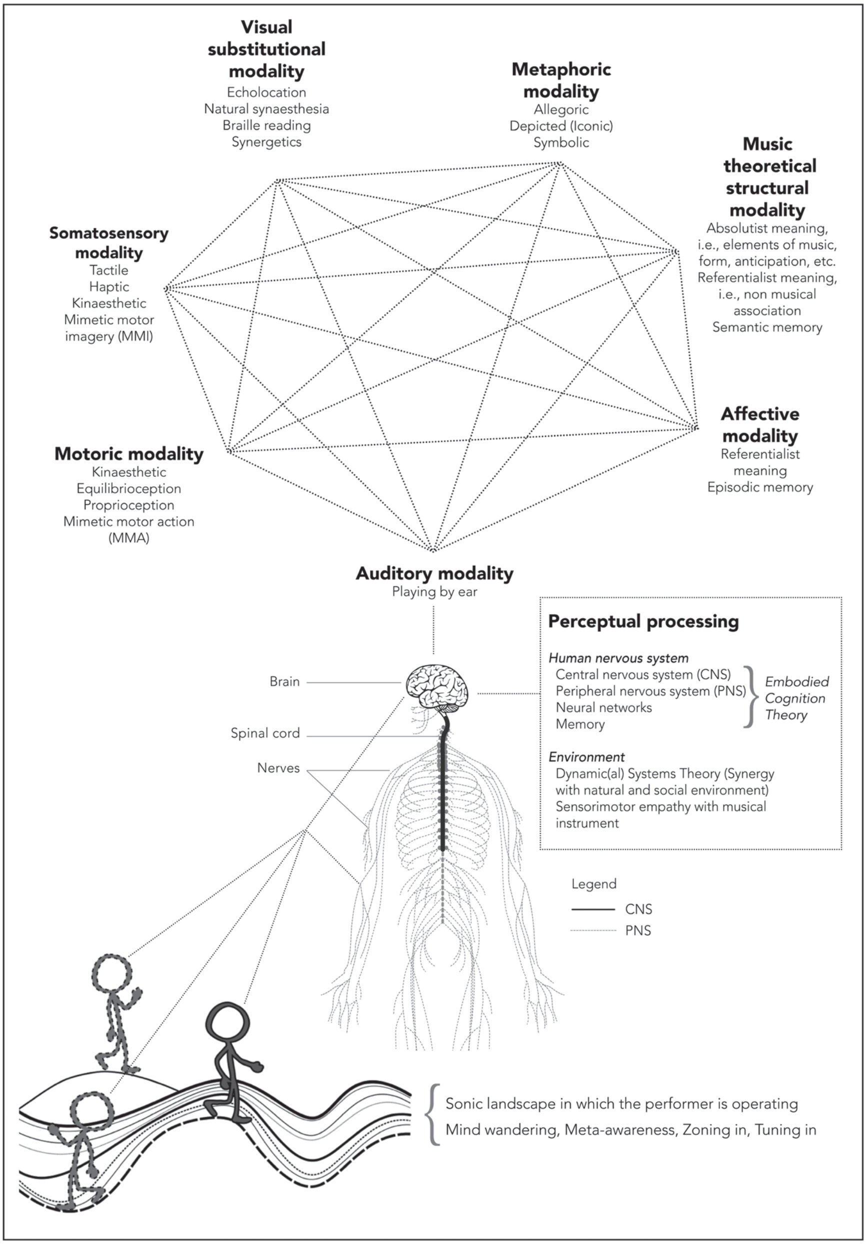


Figure 22.1 A framework of multimodal processing in visually impaired musicians such as pianists.

their instrument constitutes a cyclical process based on tactile experiences and aural feedback.

The multifaceted ideo-motor theory resurfaced together with ECT and holds that there is a link between action and perception, implying that a new modality might be formed as opposed to merely a strengthened association between the action and perception. Ideo-motor theory has a strong reliance on sensory input, which is partially inhibited in the case of VI pianists (Pfordresher, 2019).

Data analysis showed that all participants experienced different forms of mental imagery, affect, and embodied cognition within an active performance environment (DST) revealing an overlapping of finely nuanced subcategories of ECT. We concentrate on Participant D's engagement in all cognitive processes as performer and listener, as his experiences encompass and go beyond those of the other participants. He experienced synaesthesia, different forms of affect, often sublime (cf. Konečni et al., 2008); has the ability to visualize the musical score and not braille cells; could translate his mental imagery by placing himself in the role of the performer; revealed an ability to adapt to the musical performance according to the social environment (church vs. concert performance); is always involved in real-time analysis of the auditory music stimuli, considering the work as a whole; and expressed sensorimotor empathy with his instrument (synergy). He furthermore described an intimate relationship with the piano comparable to Merleau-Ponty's blind cane-user and Chemero's notion of synergetics. Ideo-motor theory could explain why he experienced problems mastering large jumps.

An Evolving Conceptual Framework

Throughout this chapter, a set of discrete yet interlinked concepts have been identified and aligned towards building a broader understanding of the symbiosis and functioning of the internal and external worlds of VI pianists (musicians). This section brings together participants' views and relevant ECT- and DST-based concepts, with reference to the literature on visual impairment and human cognition.

As a result of cross-modal plasticity, the visual sense in the VI individual is replaced with other senses (visual substitutional) such as echolocation, which uses the physics of sound for navigation; the combination of auditory and visual modalities, as found in synaesthesia; tactile sense substitution using braille in place of Western staff notation; and synergetics, which underlies an enhanced interrelationship between the performer and the instrument.

The metaphoric modality consists of the allegorical (narratives), depicted (iconic representations), and the symbolic (braille and staff notation), while concepts of absolutist and referentialist meanings of music and affect further inform this framework. All modalities could link with the affective modality in that, for example, a motoric action and referentialism could evoke an emotional response.

Sighted and VI musicians (pianists) move backwards and forwards (journeying) within a sonic landscape during learning and performing processes. Meta-awareness is highly relevant for performers who show extended levels of concentration when zoning and tuning into areas such as tone colour, form, shifting their minds temporarily and deliberately to focus on, for example, accompaniment, melody, emphasizing different parts and harmony, as well as shifting between foreground and background as explained by Gestalt theorists.

In conclusion, our investigation resulted in a multimodal framework consisting of seven inter- and cross-correlated cognition-based modalities, which function within the different levels of absorption (see Figure 22.1):

- Auditory
- Motoric
- Somatosensory
- Visual substitutional
- Metaphoric
- Music theoretical structural
- Affective

The diagram in Figure 22.1 captures perceptual processing with specific reference to VI pianists. Emphasis is placed on the cross-modality and multisensory integration implicit in ECT as well as the systemic relationships in DST. The seven modalities that cross- and interlink allow VI musicians and specifically pianists to function within a socio-musical environment. The dotted lines between the modalities indicate the systemic nature of the framework. The auditory modality forms the entry point, leading to and linking the other modalities, forming a dynamic system that is in constant flux. Modalities are organized in circular fashion to avoid suggesting a hierarchy, even though one may exist. Although the framework includes components specifically relevant to VI musicians (pianists), it could be applicable to sighted musicians if one were to exclude braille notation from the visual substitutional modality. Furthermore, one would expect a fluctuation in intensity across modalities depending on the level of visual impairment.

Conclusion

A complex set of conceptual system-related intra- and cross-cortical processes resulted into a proposed framework that used a qualitative study as a point of departure. Although participants' cumulative views provided substantial and varied perspectives, using a bigger cohort would offer the opportunity to enhance and elaborate on the emergent conceptual framework.

In repudiating the stigma associated with a disability, this study suggests that VI pianists are highly innovative in navigating obstacles.[5]

Notes

1 The World Health Organisation (WHO) estimates that about 1.3 billion people worldwide live with some form of visual impairment (www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment). A South African government census from 2011 estimated that 7,39,000 people live with a severe visual disability (www.statssa.gov.za/publications/Report-03-01-59/Report-03-01-592011.pdf).

2 Although synesthesia contains a visual element, it is a discrete modality. Also compare the chapter by Küssner and Orlandatou (this volume).
3 Three participants attended the Pioneer School for Learners with Visual Barriers previously known as the School for the Blind (est. 1881), Worcester, South Africa. One participant followed mainstream schooling, while Participant E teaches at the Pioneer school.
4 Trinity College London is a music examination board situated in the UK offering an internationally acknowledged and administered range of graded instrumental examinations (www.trinitycollege.com/resource/?id=9079).
5 This work was supported by Harry Oppenheimer Memorial Trust Study Leave Funding; University of Cape Town URC Block Grant Funding; and Master's Student Funding, University of Cape Town.

References

Andrianopoulou, M. (2019). *Aural education, reconceptualising ear training in higher music learning*. London: Routledge. https://doi.org/10.4324/9780429289767

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*(1), 1–29. https://doi.org/10.1146/annurev-psych-120710-100422

Baker, D., & Green, L. (2017). *Insights in sound: Visually impaired musicians lives and learning*. London: Routledge. https://doi.org/10.4324/9781315266060

Cattaneo, Z., & Vecchi, T. (2011). *Blind vision: The neuroscience of visual impairment*. Cambridge: MIT Press. https://doi.org/10.7551/mitpress/9780262015035.001.0001

Charmaz, K. (2014). *Constructing grounded theory* (2nd ed.). London: Sage.

Chemero, A. (2016). Sensorimotor empathy. *Journal of Consciousness Studies*, *23*(5–6), 138–152.

Covington, K. (2005). The mind's ear: I hear music and no one is performing. *College Music Symposium*, *45*, 25–41. www.jstor.org/stable/40374518

Cox, A. (2017). *Music and embodied cognition: Listening, moving, feeling, and thinking*. Bloomington: Indiana University Press.

Durie, B. (2005). Doors of perception. *New Scientist*, *185*(2484), 34–36. www.newscientist.com/article/mg18524841-600-senses-special-doors-of-perception/

Endestad, T., Godøy, R. I., Sneve, M., Hagen, T., Bochynska, A., & Laeng, B. (2020). Mental effort when playing, listening, and imagining music in one pianist's eyes and brain. *Frontiers Human Neuroscience*, *14*, 576888. http:/doi.org/10.3389/fnhum.2020.576888

Gabrielsson, A., & Lindström, E. (2010). The role of structure in the expression of emotions. In P. Juslin & J. Sloboda (Eds.), *Handbook of music and emotion: Theory and research* (pp. 367–400). Oxford: Oxford University Press.

Gallagher, S. (2014). Phenomenology and embodied cognition. In L. Shapiro (Ed.), *The Routledge handbook of embodied cognition* (pp. 9–18). London: Routledge. http:/doi.org/10.4324/9781315775845

Geeves, A., & Sutton, J. (2014). Embodied cognition, perception, and performance in music. *Empirical Musicology Review, Columbus*, *9*(3–4), 247–253. https://doi.org/10.18061/emr.v9i3-4.4538

Gordon, E. (2012). *Learning sequences in music: Skill, content, and patterns*. Chicago: GIA Publications.

Gosselin-Kessiby, N., Kalaska, L., & Messier, J. (2009). Evidence for a proprioception-based rapid on-line error correction mechanism for hand orientation during reaching movements in blind subjects. *Journal of Neuroscience*, *29*(11), 3485–3496. https://doi.org/10.1523/JNEUROSCI.2374-08.2009

Harrar, V., Aubin, S., Chebat, D.-R., Kupers, R., & Ptito, M. (2018). The multisensory blind brain. In E. Pissaloux & R. Velazquez (Eds.), *Mobility of visually impaired people* (pp. 111–136). Cham: Springer. https://doi.org/10.1007/978-3-319-54446-5_4

Herbst, A. (2021). The tortoise and the magic tree: Strategies to develop comprehensive and holistic music analytical listening skills through the use of "ghost scores". In K. Cleland & P. Fleet (Eds.), *The Routledge companion to aural skills pedagogy* (pp. 190–209). New York: Routledge.

Hiver, P., Al-Hoorie, A., & Larsen-Freeman, D. (2021). Toward a transdisciplinary integration of research purposes and methods for complex dynamic systems theory: Beyond the quantitative-qualitative divide. *International Review of Applied Linguistics in Language Teaching*. https://doi.org/10.1515/iral-2021-0022

Huron, D. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge: MIT Press. https://doi.org/10.7551/mitpress/6575.001.0001

Johnson, M. (2007). *The meaning of the body: Aesthetics of human understanding*. Chicago: Chicago University Press. https://doi.org/10.7208/chicago/9780226026992.001.0001

Johnson, M., & Larson, S. (2003). Something in the way she moves: Metaphors of musical motion. *Metaphor and Symbol*, *18*(2), 63–84. https://doi.org/10.1207/S15327868MS1802_1

Karpinski, G. (2000). *Aural skills acquisition: The development of listening, reading, and performing skills in college-level musicians*. Oxford: Oxford University Press.

Konečni, V., Brown, A., & Wanic, A. (2008). Comparative effects of music and recalled life-events on emotional state. *Psychology of Music*, *36*(3), 289–308. https://doi.org/10.1177/0305735607082621

Korsakova-Kreyn, M. (2018). Two-level model of embodied cognition in music. *Psychomusicology: Music, Mind, and Brain*, *28*(4), 240–259. https://doi.org/10.1037/pmu0000228

Kosslyn, S., Thompson, W., & Ganis, G. (2006). *The case for mental imagery*. Oxford: Oxford University Press. https://doi.10.1093/acprof:oso/9780195179088.001.0001

Lakoff, G., & Johnson, M. (2011). *Metaphors we live by: [With a new afterword]*. Chicago, IL: University of Chicago Press. https://doi.org/10.7208/chicago/9780226470993.001.0001

Laverty, S. (2003). Hermeneutic phenomenology and phenomenology: A comparison of historical and methodological considerations. *International Journal of Qualitative Methods*, *2*(3), 21–35. https://doi.org/10.1177/160940690300200303

Legge, G. E., Granquist, C., Lubet, A., Gage, R., & Xiong, Y. (2019). Preserved tactile acuity in older pianists. *Attention Perception Psychophysics*, *81*, 2619–2625. https://doi.org/10.3758/s13414-019-01844-y

Loeffler, J., Raab, M., & Cañal-Bruland, R. (2016). A lifespan perspective on embodied cognition. *Frontiers in Psychology*, *7*, 845. https://doi.org/10.3389/fpsyg.2016.00845

Loimusalo, N., & Huovinen., E. (2018). Memorizing silently to perform tonal and nontonal notated music: A mixed-methods study with pianists. *Psychomusicology: Music, Mind, and Brain*, *28*(4), 222–239. http://dx.doi.org/10.1037/pmu0000227

Mast, F., Tartaglia, E., & Herzog, M. (2012). New percepts via mental imagery? *Frontiers in Psychology*, *3*, 360. https://doi.org/10.3389/fpsyg.2012.00360

Matya, J. (2016). Embodied music cognition: Trouble ahead, trouble behind. *Frontiers in Psychology*, *7*, 854. https://doi.org/10.3389/fpsyg.2016.01891

Meyer, L. (1956). *Music and meaning*. Chicago, IL: Chicago University Press.

Moulton, S., & Kosslyn, S. (2009). Imagining predictions: Mental imagery as mental emulation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(521), 1273–1280. https://doi.org/10.1098/rstb.2008.0314

Nanay, B. (2018). Multimodal mental imagery. *Cortex*, *105*, 125–134. https://doi:10.1016/j.cortex.2017.07.006

Ockelford, A., & Matawa, C. (2009). *Focus on music 2: Exploring the musicality of children and young people with retinopathy of prematurity*. London: Institute of Education.

Park, H., & Kim, M. (2014). Affordance of Braille music as a mediational means: Significance and limitations. *British Journal of Music Education*, *31*(2), 137–155. https://doi.org/10.1017/S0265051714000138

Parreira, B., Grecco, L., & Oliveira, O. (2017). Postural control in blind individuals: A systematic review. *Gait and Posture*, *57*, 161–167. http://dx.doi.org/10.1016/j.gaitpost.2017.06.008

Patestas, N., & Gartner, L. (2016). *A textbook of neuroanatomy* (2nd ed.). Hoboken: Wiley.

Pearson, J., Naselaris, T., Holmes, E., & Kosslyn, S. (2015). Mental imagery: Functional mechanisms and clinical applications. *Trends in Cognitive Sciences*, *19*(10), 590–602. https://doi.org/10.1016/j.tics.2015.08.003

Perlovsky, L. (2015). Origin of music and embodied cognition. *Frontiers in Psychology*, *6*, 538. https://doi.org/10.3389/fpsyg.2015.00538

Perone, S., & Simmering, V. (2017). Applications of dynamic systems theory to cognition and development: New frontiers. In J. Benson (Ed.), *Advances in child development and behaviour* (Vol. 52, pp. 43–80). London: Elsevier. https://doi.org/10.1016/bs.acdb.2016.10.002

Petrović, M., & Golubović, M. (2018). The use of metaphorical musical terminology for verbal description of music. *Rasprave*, *44*(2), 627–641. https://doi.org/10.31724/rihjj.44.2.20

Pfordresher, P. (2019). *Sound and action in music performance*. London: Academic Press.

Schenkman, B., & Nilsson, E. (2010). Human echolocation: Blind and sighted persons "ability to detect sounds recorded in the presence of a reflecting object". *Perception*, *39*(4), 483–501. https://doi.org/10.1068/p6473

Schmidt, T., & Blankenburg, F. (2019). The somatotopy of mental tactile imagery. *Frontiers in Human Neuroscience*, *13*, 10. https://doi.org/10.3389/fnhum.2019.00010

Shapiro, L. (2019). *Embodied cognition* (2nd ed.). New York: Routledge. https://doi.org/10.4324/9781315180380

Taruffi, L., & Küssner, M. (2019). A review of music-evoked visual mental imagery: Conceptual issues, relation to emotion, and functional outcome. *Psychomusicology: Music, Mind, and Brain*, *29*(2–3), 62–74. https://doi.org/10.1037/pmu0000226

Thelen, E. (2005). Dynamic systems theory and the complexity of change. *Psychoanalytic Dialogues*, *15*(2), 255–283. https://doi.org/10.1080/10481881509348831

Van der Schyff, D., Schiavio, A., Walton, A., Velardo, V., & Chemero, A. (2018). Musical creativity and the embodied mind: Exploring the possibilities of 4E cognition and dynamical systems theory. *Music and Science*, *1*, 205920431879231. https://doi.org/10.1177/2059204318792319

Vroegh, T. (2019). Zoning-in or tuning-in? Identifying distinct absorption states in response to music. *Psychomusicology: Music, Mind, and Brain*, *29*(2–3), 156–170. https://doi.org/10.1037/pmu0000241

Wootton, J. (2005). *Teaching braille music notation to blind learners using the recorder as an instrument* [Doctoral dissertation, Stellenbosch University]. SUNScholar Research Repository. http://scholar.sun.ac.za/handle/10019.1/50461

23 "Don't Sing It With the Face of a Dead Fish!" Can Verbalized Imagery Stimulate a Vocal Response in Choral Rehearsals?

Mary T. Black

Imagery has been verbalized by choral directors in rehearsals for over a century (Coward, 1914), but which modes of *verbalized imagery* (VI) are evident in this context? Are there particular types of vocal response for which VI is most frequently used? More importantly, can VI affect directors' practice? Experience as singer and choral director prompted those questions; this chapter seeks to answer them, beginning with a definition of VI in the context of choral rehearsals. The chapter demonstrates that directors employ VI to create perceptible changes in vocal responses.

In choral rehearsals, imagery is presented verbally from director to choir members, with the intention of affecting the vocal response. The VI is interpreted by the singers who create or change their vocal response. Choral director Scott provided the title phrase which provoked his singers to mentally envisage the *face of a dead fish* and how their faces might be resembling it, for example, with the mouth turned down at the corners. As the VI was humorous, singers smiled in response which changed their facial expressions with the result that it immediately altered the tone quality. This example (see Table 23.1, number 1) illustrates a typical interaction where VI is employed to guide the vocal responses of choir members.

Jacobsen defines verbal imagery as "mental pictures created through spoken figures of speech" (2004, p. 3). In this chapter, the term *verbalized imagery* is being used: this is an image, metaphor, simile, or other figurative language which is employed verbally by choral directors and heard by singers, the purpose of which is to enhance directors' explanations in order to affect singers' vocal responses in rehearsals. This definition employs VI as a collective noun for a range of literary devices, as does Spitzer (2004); elsewhere, Ortony notes the potential of figurative language to "transfer learning and understanding" (1975, p. 53). The definition also highlights the context of regular choral rehearsals.

While seeking the aforementioned definition, it became apparent that VI is not the same as *auditory verbal imagery (Shergill et al., 2001)*, where the speech or voice is to be imagined. Neither is it *musical imagery* (Godøy & Jørgenson, 2001), where again the musical sound is absent, to be imagined rather than produced

DOI: 10.4324/9780429330070-28

(Floridou, this volume). In choral rehearsals, singers actually hear VI and its purpose is for them to vocally produce the sound. The VI is the trigger and the sound is the result; VI is used to help the singers imagine something that supports activation of part of their physical apparatus to respond in a way which produces the sound the director requires. The VI in choral rehearsals is therefore applied by directors to elicit singers' thoughts about the VI which prompts perceptible vocal effects (VEf).

Regarding the modes of VI (i.e., the mode of the VI example), decades of vocal tuition (Williams, 2013) demonstrate that the connection between VI and visualized images is accepted and includes the notion of pictures in the mind; visual and emotional modes were found to be important in creating the VEf (Callaghan et al., 2012; DeGroot, 2009; Vesely, 2007; Williams, 2013), as music can relate to and release the emotions (Emmons & Chase, 2006). In relation to the type of VEf produced, several sources (e.g., Decker & Kirk, 1988; Miller, 1996), suggest that VI is useful only for altering expression and feelings, which might imply these VEf are more ephemeral. However, Hemsley disagrees saying "singers should not learn a piece and then attempt to add the feeling and expression afterwards" (1998, pp. 111–112), indicating the more crucial role of feelings and expression. The "inexpressibility" of VI which allows the "transfer of characteristics which are unnameable" (Ortony, 1975, p. 48) might be particularly useful for describing tone quality and expression, especially with amateur singers. One of Durrant's requirements for choral directors is "the capacity to communicate clearly and unambiguously" (2003, p. 100). Jansson too requires "sensemaking" to be a "function of the conductor's position" (2014, p. 151), which prompted the question whether VI could positively affect directors' practice.

In order that any results of the research could be applicable to real-life rehearsals, data were sought concerning the role of VI, particularly in relation to any changes in VEf. One of the initial ideas to be explored was that VI explains what cannot be seen. First, this relates to the inner workings of the singing mechanism; much vocal pedagogy relies on singers having a conceptual understanding of physiology, some of which is not visible. Zbikowski's conceptual models relating to the voice are essential here, as "it is difficult to find the physical object that correlates with vocal sound" (2002, p. 111), which is a particular problem for amateur singers. The second invisible aspect is what does not appear in the notation but is heard in performance, that is, the nuances of the sound which might be termed expression or interpretation. In the same way that a composer "articulates subtle complexes of feeling that language cannot even name" (Langer, 1969, p. 222), singers need to interpret these. Schippers too noted the need for something "beyond the tangible" (2006, p. 214). Some aspects of expression and interpretation do appear in the score and choral directors may explain this verbally. However, choirs who use notation need, at some point, to focus more on the sound produced than on the written page. The notation guides singers and directors but there is much of the overall performance which is not in the notation. It might be in cases like these that VI is most useful.

Method

The data for this qualitative research were collected over a period of five years and completed in two phases: the first to test validity and reliability for the second.

Participants

Overall, 15 amateur choirs (332 singers) and 21 directors, some professional, contributed to the research. All respondents were given pseudonyms[1] and were told the research was investigating rehearsal strategies[2] to avoid bias in their responses. Five directors, Sam, Pete, Emma, Ken, and Tim and their choirs, participated in four types of materials, for example, questionnaires, interviews, observations, and videoed rehearsals. This approach produced the most comprehensive and reliable data and therefore provided the majority of results.

Materials and Procedures

Observations and video recordings captured the VI in regular rehearsals rather than experimental situations; obtaining permissions for these restricted the number of recordings made though they provided the most reliable data in terms of capturing VI. During the interviews, singers and directors (separately) viewed excerpts of their previous week's rehearsal at points where the choir had sung a phrase or note, the director had used VI and the choir had sung the same phrase again. Respondents were asked to focus on and compare the vocal responses pre- and post-VI to explain whether they heard any differences and if so, what those were. During regular rehearsals, the director judges whether vocal responses are appropriate; therefore, their reactions to reviewing the responses during the interviews were deemed key to deciding what VEf[3] were heard and whether their intentions had been achieved, for example, whether the tone quality had changed.

Analysis

Interpretative Phenomenological Analysis (Smith et al., 2009) was employed as the most appropriate analytical strategy due to its focus on how participants make sense of their experiences. The main themes to emerge from the analysis which are relevant to this chapter were the modes of VI and types of VEf. Data reduction was based on relevance to the themes, for example, how the vocal response was affected, rather than prevalence, though the number of comments per theme was also logged.

The VI gathered in Phase 1 of the research was classified according to six modes: Visual, Auditory, Kinaesthetic, Emotional, Conceptual—Musical, and Conceptual—Non-musical. The first three modes correspond with the modes of presentation of learning Fleming (1995) devised. Other initial ideas for the modes, such as the inclusion of emotional imagery, were taken from poetic imagery and figurative language, to which Cuddon adds that feelings, thoughts, states of mind and any sensory or extra-sensory experiences may be denoted by imagery (Cuddon, 1979). These ideas widen the range of modes considerably and partly indicate why VI is so prevalent. In addition, Paivio and Begg (1981) include conceptual images, which were sub-divided into conceptual—non-musical (e.g., Table 23.1, number 2 *splendour*), and conceptual-musical modes (see Table 23.1, number 3 *trumpet*). The inclusion of conceptual modes signalled that thoughtful reflection on the VI was involved in the process of singers deciding how to create the vocal response.

Multimodal VI (MmVI), that is, VI which exists in more than one mode (Algozzine & Douville, 2004; Mountain, 2001) is also in evidence, for example, Table 23.1, number 4 *jewel*, which is a combination of conceptual—non-musical and visual modes. MmVI allows directors to access multiple learning styles simultaneously (Durrant, 2003), and to encourage the inclusion of the imagination (Williams, 2013). The close interrelationship between the modes in MmVI mirrors that between the different elements of vocal technique including interpretation, expression, and accuracy. MmVI therefore frees any constraints of separation between those elements so becomes invaluable to directors as they can affect several VEf simultaneously.

In addition to the modes of VI, a set of categories was devised to classify the types of VEf (see Table 23.1, Note +). The main categories were Voice Production and Technique, Expression/Interpretation, Motivation, and Musical Elements. The categories are focused on the *effect* of the VI, that is, any change in vocal response which the singers made to the VI.

The categories were informed by several authors, including Chen (2007), whose research focused on verbal imagery and individual singers, and Jacobsen (2004), who analysed imagery in choir settings but sought information only on vocal function. The categories are sub-divided and cover a wide range of VEf, for example, tone quality and expression.

Results and Discussion

This section will focus on two of the modes and one of the roles of VI.

Modes

The frequency counts and percentages of modes of VI of Phase 1 directors whose rehearsals were videoed (see Table 23.2, Supplementary Material) show that nearly all examples (i.e., 116 out of 136) were multimodal. For example, all of Rob's modes which were visual, that is, 3, were also multimodal, though only 13 of the 15 conceptual-musical modes were multimodal. Although Rob provided 53 VI examples in total, they exhibit 93 modes, as most VI examples were multimodal.

The largest proportion of modes for all directors is conceptual, that is, 46.6% of those provided by Sam and more than 60% of each of Rob and Pete (see Table 23.2, Supplementary Material). Of those, the non-musical concepts outweigh the musical concepts though only slightly in the case of Sam. Imagery (verbalized in this context) is extremely useful for developing conceptual understanding (Ortony, 1975; Paivio & Begg, 1981), especially in relation to some of the invisible aspects of singing mentioned earlier, for example, tone quality.

Visual and Emotional Modes

The importance of visual and emotional modes (Callaghan et al., 2012; Williams, 2013) was noted earlier, and since relatively few were detected in the research, it is interesting to examine some which were located.

Pete's singers were imagining the sight of the *jewel*, for example, how brightly it sparkled (see Table 23.1, number 4). He made a direct comparison between the note and the idea of shining *through* the vocal texture; therefore, the modes are visual and conceptual—non-musical.

The sight of a glittering Viennese ballroom, for example, with sparkling chandeliers and ornate paintings (see Table 23.1, number 5) is also visual. However, singers would need to understand what constituted that *Viennese glitter* before they were able to create a visual image. The *Viennese glitter* in example 5 combines visual and conceptual—non-musical modes with the VEf being a change in tone quality and expression to provide the additional *sparkle*.[4]

The second mode to be examined is the emotional mode. Rob asked for an emotional response using the idea of *chuckling* (see Table 23.1, number 6), referring to the concept rather than suggesting singers actually chuckled while singing. The phrase sounded more rhythmically articulated, separating the sounds in imitation of the chuckle.

Rob highlighted what he wanted to eradicate in the next example (see Table 23.1, number 7). He required the singers to imagine the emotion of being *startled* but to replace it with the opposite idea by preparing their breath with confidence. Singers were motivated to start the note with a more precise phonation and continue with consistent breath flow.

Further Examination of Modes

The high percentage of conceptual as opposed to other modes of VI seems to indicate a more complex view of VI than simply the illustration of an idea with a single sensory example. If this were the case, more visual, aural, and kinaesthetic modes might be expected. That complexity is mirrored by the variety of VEf which were produced. Given the connection between VI and visualized images (Jacobsen, 2004), the small percentages of visual modes (see Table 23.2, Supplementary Material), provided by Rob and Pete are unexpected. It is also worth noting that among these data, none of the visual modes envisaged the inner physiological processes of singing. With these directors at least, visualizing how the sound is created in the body was not a priority, as their concern was with the VEf rather than with using VI to obtain it; the VI is a means to an end.

In contrast to findings noted earlier (e.g., Emmons and Chase, 2006) none of the directors used the emotional mode more than modes related to the senses. There

Table 23.1 Verbalized imagery examples with director, mode, and vocal effect.

Examples	*Director*	*Mode#*	*Vocal Effect+*
1. Don't sing it with the face of a dead fish!	Scott	Vis	Po/T
2. Piling on the splendour all the time.	Rob	Cn	EX
3. Like a trumpet!	Pete	Au/Cm	T
4. I want that C# to shine through. I want it to be like a jewel suddenly being caught by the light.	Pete	Vis/Cn	T/TX
5. It's not got the Viennese glitter.	Rob	Vis/Cn	T/EX
6. Don't forget we chuckle there tenors and basses.	Rob	E/Cn	R/A
7. You sound like you're being startled!	Rob	E	M/B/V
8. You [tenor] can have a bit more stability in your tone.	Emma	Cn	T
9. Make that really strong, you know . . . very . . . bell-like . . . don't hold back.	Ken	Cn/Cm	T
10. So now you know what you're going to sing, let's get a proper *Hallelujah* this time.	Tim	Cm	EX/M
11. *Trumpets, clanger, excites*, can you do something with those words . . . a little reminder to bring it alive.	Ken	Cn	EX/A

Note: # Key to Modes

Vis—Visual; Au—Aural; K—Kinaesthetic; E—Emotional; Cm—Conceptual: musical; Cn—Conceptual: non-musical.

Note: + Key to Vocal Effects

B—Breath management/control, support, respiration, energy; T—Tone quality, register, tone colour, resonance, vibrato; V—Voice production, projection, how or where sound is created, larynx and phonation; Po—Posture, stance, body position/alignment, facial expression, mouth shape; A—Articulation consonants, vowel shape or formation, diction, pronunciation (including glottal stops); F—Flow of piece/phrase; line, shape of phrase, phrasing, how phrase moves, urgency (but not where this is related to text, for example, commas and not flow of breath); EX—Expression, interpretation, imagination, may or may not refer to text; M—Motivation, enthusiasm, readiness, concentration, confidence, alertness, use of humour to motivate; D—Dynamic, volume, emphasis, accent, stress; R—Rhythm, rhythmic reading and accuracy, detached, staccato, timing, entries together, ensemble; Pi—Pitch, intonation, range, melodic reading and accuracy; TX—Texture, balance of voices, parts interweaving, one voice part standing out; S—Speed, tempo, metronomic pulse keeping.

is no emerging pattern of which modes are most often used overall by Rob, Sam, and Pete. Directors were not asked about their preference for particular modes during data collection in order not to prejudice the outcomes, so although the lack of any pattern here may be due to their personal preferences, the data are necessarily inconclusive.

Conceptual modes were the most frequently encountered (see Table 23.2, Supplementary Material), but of the others (visual, auditory, kinaesthetic, and emotional), no pattern emerged of those used most often. The majority of VI examples were multimodal though there were 20 conceptual (musical and non-musical) single-mode VI (marked [a, b] and [c] in Table 23.2, Supplementary Material), 17 from Rob and 3 from Pete. However, there was insufficient evidence to link those 20 single modes with a particular vocal effect (see Table 23.3, Supplementary Material). In fact, half the conceptual single-mode VI produced multiple VEf thereby impairing the possibility of determining whether there was a connection between, for example, the visual mode and breath management, or between the emotional mode and motivation.

In addition, the current research was initially intended to be applicable to regular choral rehearsals in terms of the VEf of the VI in singers' vocal responses. As the relationship between specific modes and particular VEf was inconclusive, the decision was made to discontinue examining the modes in the second phase of the research and focus primarily on the VEf of the VI, which could be beneficial to directors in terms of altering singers' vocal responses overall.

One Role of Verbalized Imagery: VI Can Produce a Change in Vocal Effect

The most important of the nine roles[5] of VI which were found during the research as a whole was that it was able to produce a change in VEf. If VI was not functioning in that way, it would be of little use to directors; for that reason it has been chosen as a focus for this chapter.

During the interviews, directors listened intently to the responses singers made after VI had been employed. They decided that the responses had changed, and their intentions had been fulfilled, while acknowledging a few cases where the sound had altered insufficiently, or where their VI was misleading so the response did not adequately match their intentions. In those cases, they modified the VI or changed strategy accordingly. During their rehearsals, directors were not consciously choosing specific VI, or modes of VI but spontaneously seeking the most appropriate vocabulary for their explanations. Singer responses were affirmative not only that the sound had changed but also, with one exception, that the director's requirements had been achieved.

A surprising discovery, in contrast to research noted earlier (e.g., Miller, 1996) was that VI was most frequently employed to affect tone quality (16.61%) rather than expression (9.71%) (see Figure 23.1). The frequency counts and percentages of each of the resulting VEf (see Table 23.4, Supplementary Material, and Figure 23.1) also show some individual differences between those directors whose rehearsals were videoed.

Tone Quality and Expression

It is worthwhile examining examples of tone quality and expression, given VI's potential for inexpressibility noted earlier.

The VI *stability* and *strong* (see Table 23.1, numbers 8 and 9) are similar in their outcome. The tone quality is strengthened by creating more consistent breath

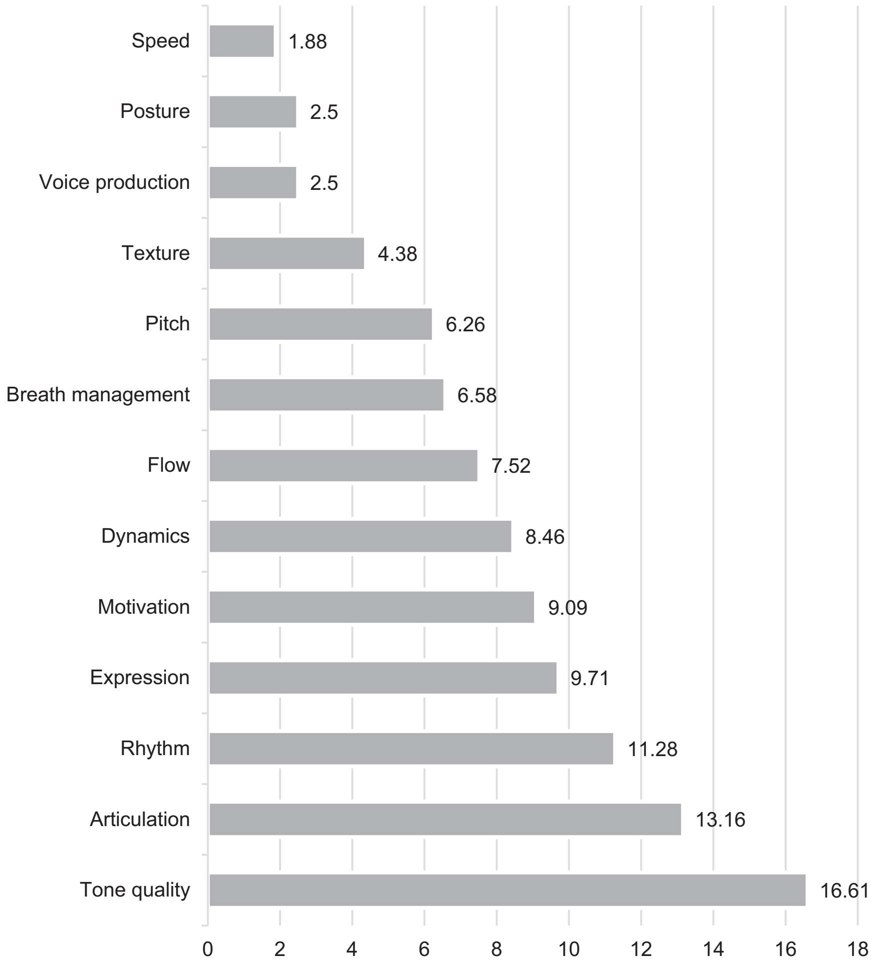


Figure 23.1 A bar graph to show percentages of vocal effects of verbalized imagery across directors.

flow, keeping the back of the throat open, and raising the soft palate; this would also supply the more ringing tonal quality which Ken additionally required.

The *Hallelujah* Tim referred to (see Table 23.1, number 10) is not in the text, but it is a musical idea with which his gospel choir were familiar, music of lively praise and joy. Once the singers were familiar with the words, rhythms, and pitches, he then employed VI to induce them to concentrate more on the expression of praise and joy. Tim taught his singers by rote and this may explain why their rehearsal process conflicts with Hemsley's quote, earlier.

Ken created the VI with three words from the text which focus on a call to battle, *trumpets, clanger* and *excites* (see Table 23.1, number 11). The words summarized the ideas of the whole movement as Ken tried to generate in the singers' expressions of bringing the sound "alive," which would be heightened by clearer articulation of consonants as they would be more noticeable and so liven up that section.

In relation to the VEf of VI, the fact that tone quality was the most frequently observed vocal effect overall is very important to directors firstly, because it is often difficult to describe (Ortony, 1975), especially to amateur singers. Directors were employing VI to avoid technical terminology or provide more detail or a different portrayal of what singers were aiming towards for a particular vocal effect. Secondly, in addition to working on note-accuracy, directors can focus on improving the standard of their choir's singing by balancing and moulding the overall tone quality. Concentration on this, often termed *choral blend* (Chipman et al., 2008) was the main result of Emma's VI for her singers (see Table 23.4, Supplementary Material), who were the most vocally skilled choir in the research.

As stated earlier, some parts of the vocal apparatus are unseen and in addition "the mechanism involved is complex, variable, and never entirely under conscious control" (Bartholomew, 1940, p. 20). It is therefore essential for directors to be able to create and change sounds by affecting what singers are thinking. This is possibly the reason for the higher proportions of conceptual VI and of VI affecting tone quality. The VI is being employed to build conceptual understanding, particularly for concepts which are abstract, or not under the singers' direct control; this is important to directors. It should not be assumed that VI only allows singers to imagine seeing a visual image. It also enables them to think of the VI, to have some understanding of what makes, for example, the *jewel* stand out from its background, and to respond vocally to it, all within milliseconds during a rehearsal.

Several of the examples in Table 23.1 were multimodal and produced multiple VEf, for example, number 4. A pattern found throughout the research was that most VI produced multiple VEf. This result might be due partly because the VEf are so closely related, partly because the music itself is difficult to compartmentalize and partly because of the holistic nature of VI which is so full of meaning; this is discussed at some length elsewhere (Black, 2015, pp. 160–163), as is some of the multieffect VI.

Conclusion

This chapter has demonstrated that it is possible to categorize the modes of VI directors use in their rehearsals, though there is no evidence to suggest a link between a particular mode and a specific vocal effect. This is mainly because of the plethora of MmVI which produced multiple VEf in singer responses.

More importantly, evidence perceived by directors has shown that VI produced intended changes in VEf. The VI enabled directors to refer to invisible aspects of the singing mechanism. The VEf were wide-ranging and included some which were voice-specific, for example, tone quality, and some which were not, for example, expression. A change in tonal quality was the vocal effect encountered most frequently with VI and it would be interesting to investigate if this were replicated in future studies with different participants. Research might also be devised comparing the VEf of VI in amateur and professional settings.

The findings presented in this chapter should encourage directors to employ VI in their rehearsals alongside other strategies, such as the vocal demonstrations or gestures listed earlier. Directors need not be sceptical about VI, regarding it as vague or only relevant to expression or interpretation. Imagery forms an integral part of everyday speech (Paivio & Begg, 1981, p. 287), particularly in verbal explanations; therefore, it is not difficult to incorporate VI into rehearsals. The knowledge that VI enabled specified VEf to be produced should persuade directors to apply it with confidence, as a valuable and beneficial strategy in their choral rehearsals.

Notes

1 Pseudonyms used throughout.
2 This would include vocal demonstrations, gestures, and technical terminology, though these were not exclusive and were not specified to respondents.
3 See Table 23.1, Note + categories of vocal effect.
4 In this chapter, it is not possible to hear the musical phrases to which these examples relate; therefore, it is more difficult to comprehend their VEf.
5 Other roles focused on the use of VI as a mnemonic, transmitting and achieving director objectives, changing thinking, creating multiple-effects, saving rehearsal time, replacing technical terminology, illustrating the text and association with a specified musical phrase. These were generated from the whole research and are not evidenced in the current chapter.

References

Algozzine, B., & Douville, P. (2004). Use mental imagery across the curriculum: Tips for teaching. *Preventing School Failure: Alternative Education for Children and Youth*, *49*(1), 36–39. doi:10.3200/PSFL.49.1.36

Bartholomew, W. T. (1940). The paradox of voice teaching. In W. T. Bartholomew (Ed.), *The role of imagery in voice reaching* (1951, Reprinted from the *Journal of the Acoustical Society of America*, Vol 11, No. 4, 446–450, April 1940 ed., pp. 20–24). Jacksonville, FL: National Association of Teachers of Singing. doi:10.1121/1.1916058

Black, M. T. (2015). *Let the music dance! The functions and effects of verbal imagery in choral rehearsals* [Doctoral Dissertation, University of Leeds]. etheses. http://etheses.whiterose.ac.uk/id/eprint/13121

Callaghan, J., Emmons, S., & Popeil, L. (2012). Solo voice pedagogy. In G. McPherson & G. T. Welch (Eds.), *The Oxford handbook of music education* (Vol. 1, pp. 559–580). New York: Oxford University Press.

Chen, T.-W. (2007). *Role and efficacy of verbal imagery in the teaching of singing*. Saarbrücken: VDM Verlag Dr. Müller.

Chipman, B. J., Hoffman, J., & Thomas, S. (2008). *Singing with body, mind and soul: A practical guide for singers and teachers of singing*. Tucson: Wheatmark.

Coward, H. (1914). *Choral technique and interpretation*. London: Novello.

Cuddon, J. A. (1979). *A dictionary of literary terms*. London: Penguin.

Decker, H. A., & Kirk, C. J. (1988). *Choral conducting: Focus on communication*. Long Grove: Waveland Press Inc.

DeGroot, J. (2009). Chorus and vocal: Steps towards "visualizing" in vocal practice. *Teaching Music*, *16*(6), 63.

Durrant, C. (2003). *Choral conducting: Philosophy and practice*. New York: Routledge Taylor and Francis.

Emmons, S., & Chase, C. (2006). *Prescriptions for choral excellence*. New York: Oxford University Press.

Fleming, N. D. (1995). I'm different, not dumb: Modes of presentation (VARK) in the tertiary classroom. In A. Zelmer (Ed.), *Proceedings of the 1995 Annual Conference of the Higher Education and Research Development Society of Australasia* (Vol. 18, pp. 308–313). Canterbury, New Zealand: Higher Education and Research Development Society of Australasia.

Godøy, R. I., & Jørgenson, H. (2001). *Musical imagery*. Lisse: Swets and Zeitlinger BV.

Hemsley, T. (1998). *Singing and imagination*. New York: Oxford University Press.

Jacobsen, L. L. (2004). *Verbal imagery used in rehearsals by experienced high school choral directors: An investigation into types and intent of use* [Unpublished DMA Thesis]. University of Oregon.

Jansson, D. (2014). Modelling choral leadership. In U. Geisler & K. Johansson (Eds.), *Choral singing* (pp. 142–164). Newcastle upon Tyne: Cambridge Scholars Publishing.

Langer, S. (1969). *Philosophy in a new key* (3rd ed.). London: Oxford University Press.

Miller, R. (1996). *On the art of singing*. New York: Oxford University Press.

Mountain, R. (2001). Composers and imagery: Myths and realities. In R. I. Godøy & H. Jørgenson (Eds.), *Musical imagery*. Lisse: Swets and Zeitlinger BV.

Ortony, A. (1975). Why metaphors are necessary and not just nice. *Educational Theory*, *25*(1), 45–53. https://doi.org/10.1111/j.1741-5446.1975.tb00666.x

Paivio, A., & Begg, I. (1981). *Psychology of language*. Englewood Cliffs: Prentice-Hall Inc.

Schippers, H. (2006). "As if a little bird is sitting on your finger . . .": Metaphor as a key instrument in training professional musicians. *International Journal of Music Education*, *24*(3), 209–217. https://doi.org/10.1177%2F0255761406069640

Shergill, S., Bullmore, E., Brammer, M., Williams, S., Murray, R., & McGuire, P. (2001). A functional study of auditory verbal imagery. *Psychological Medicine*, *31*(2), 241–253. https://doi.org/10.1017/S003329170100335X

Smith, J. A., Flowers, P., & Larkin, M. (2009). *Interpretative phenomenological analysis*. London: Sage Publications.

Spitzer, M. (2004). *Metaphor and musical thought*. Chicago: Chicago University Press.

Vesely, P. (2007, Summer). Teaching visual imagery for vocabulary learning. *Academic Exchange Quarterly*, *11*(2), 51–55.

Williams, J. (2013). *Teaching singing to children and young adults*. Oxford: Compton Publishing.

Zbikowski, L. (2002). *Conceptualising music: Cognitive structure, theory and analysis*. New York: Oxford University Press.

Part V

Outlook

24 Future Perspectives and Challenges

Tuomas Eerola

Music and mental imagery has clearly gained status as an emerging topic within music research, evidenced by the number of studies, special issues, and this volume of chapters. The reason for the steady rise in popularity of the broad theme of music and mental imagery, including auditory, visual, kinaesthetic, and motor imagery to name the most common ones, is easy to understand, since the topic captures the interest of a range of disciplines (e.g., music psychology, musicology, neuroscience, music education, and music therapy) and lends itself to a wide variety of approaches (from first-person phenomenology to surveys to neuroimaging studies). But as with all broad and appealing topics, this can lead to a fractionated and complex situation, where even the basic definitions are not shared by the scholars tackling the theme. I will briefly address the issues of defining and capturing mental imagery in its different varieties (e.g., visual, auditory, kinaesthetic, and multimodal) in this chapter, and outline perspectives, particularly that of embodied cognition, that might be helpful in bringing some separate threads together and prioritize the list of challenges that need to be tackled before a coherent research programme is articulated. Although the scope of attention in mental imagery related to music spans auditory, visual, kinaesthetic, and motor imagery, I will initially review mental imagery with respect to music broadly and focus on auditory imagery and music-related visual imagery as the themes that have gained most interest in this topic previously, which is also reflected in the chapters of this volume.

Over the course of music and mental imagery research, there has been a consistent emphasis on mental imagery as:

- heterogeneous, multimodal experiences, which are defined as imagining sensory objects without the actual sensory stimulus;
- phenomena that can link music with the phenomenal qualities of experiences (emotions, dreaming, mind-wandering, thinking, and altered states of consciousness);
- contextual in nature with a diverse range of triggers and generators (music, memory, situation, volition, etc.);
- a topic that needs to adopt methodologies from research on consciousness and attention;

DOI: 10.4324/9780429330070-30

- a topic that is inherently multidisciplinary and interdisciplinary;
- a useful concept that has provided terminology and tools to shape music performance, music education, and music therapy.

In this chapter, I will first focus on definitions of imagery, then move on to the issues of assessing and capturing imagery, discuss the possibility of adopting an embodied cognition perspective to incorporate multimodality naturally into the topic, and finally summarize the short-term challenges ahead. With respect to the chapters of this book, I will not attempt to systematically refer to each chapter, although I will rely on certain chapters as key summaries of an emerging perspective that is essential to discuss the next steps.

Defining Imagery: Theoretical Challenges

Music-related mental imagery research itself has not been very active in proposing new theories of its own, since it has inherited a host of eminent theories from other mental imagery research, mainly from psychology and linguistics (Kosslyn et al., 2006; Pylyshyn, 1973), as well as research on memory (Baddeley & Andrade, 2000) and learning (Tartaglia et al., 2009). It seems that the challenges are not so much in the lack of theories overall but that these theories are usually developed within each modality or specialist topic area. Moreover, the trend of connecting specific aspects of mental imagery into broader cognitive processing such as mind-wandering, attention, or multimodal and cross-modal processing seems to be challenging (Nanay, this volume). There is some theorizing across the modalities, and areas such as the chapter by Godøy (this volume) postulate that mental imagery is a simulation of the real world where mental imagery resembles real-world sensations and the act of simulation is covered by motor mimetic cognition (Godøy, 2001, 2003). His chapter puts forward a fascinating and articulated proposal of how complex and continuous musical actions can be implemented by the human mind in an ecological fashion. The answer lies in saving up processing by controlling the actions intermittently, which connects the actions to motor patterns and overcomes the slow speed of our reactions. This is perhaps geared initially towards imagining sounds/music and not so much towards mental imagery of other modalities. But it does outline more broadly how musical imagery may trigger motor cognition and how the processes involved in this complex cognitive process could operate.

If we constrain the discussions to musical imagery, the range of issues to be mindful in theorizing and planning empirical studies is inconveniently wide. For instance, for musical imagery, the features of interest can vary along multiple features: pitch, contour, loudness, tonality, timbre, and lyrics to name a few. And when you add the different modalities (kinaesthetic, spatial, visual, motor, see Nanay, this volume, for a full overview) as well as delineating the voluntary/involuntary nature of the phenomenon, we are not talking about simple theories with multiple, interchangeable components anymore but fundamentals of how the human mind creates meaning and interprets internal states and the external world. In this respect, many authors in this volume create boundaries that constrain the scope of music-related mental imagery; Floridou (this volume) spells out a useful division of music-related mental imagery that relates to the presence of music in mental imagery, specifying that mental imagery can occur *before* (voluntary and novel mental imagery), *during* (constructive and anticipatory imagery), and *after* playing/listening (earworms, everyday visual mental imagery, etc.). Most of these three sequences do get detailed attention among the chapters such as mind-wandering (Konishi, this volume), earworms (Floridou, this volume) or altered states of consciousness (Fachner, this volume). Another interesting evaluation of the scope of music-related mental imagery is carried out by Jakubowski (this volume). She focuses on music-evoked autobiographical memories (MEAMs), where the dilemma is that it is difficult to know how much the encoding of the memories are visual, semantic or even kinaesthetic or motoric. Her findings suggest that the memories cued by music have significant embodied components and also that visuo-spatial and motor imagery are not uncommon in the descriptions of the MEAMs. This should not really be surprising as experiencing music (performing and listening) is usually a multimodal experience, but it again reminds us to include different modalities or the broader theoretical framework in future theorizing and empirical work about music and mental imagery.

Assessing Imagery: Empirical Challenges

It is crucial that we are able to capture musical imagery and other types of mental imagery related to music-making and listening in a way that is precise, both in qualitative terms but also as a temporal process. These are both challenging tasks since much of this topic consists of fleeting imagery (of various modalities) that is subject to attentional fluctuations which are heavily influenced by turning one's attention to them, as usually happens with all self-report methods. And as the episodes of music-related mental imagery are captured, the nuances need to be preserved such as the strength or vividness of the imagery, and the actual imagery details in whatever modality is in question.

Diverse sets of methods in empirical research have been wielded towards music-related mental imagery; from first-person accounts to structured interviews and surveys to self-report evaluation tasks in experiments using rating scales as well as reaction time tasks (for a full account, see Hubbard, this volume). All self-report measures have shortcomings in reliability and consistency, and Hubbard suggests that non-verbal and indirect measures (production tasks such as drawing lines according to the melodic contour, or tapping according to the tempo of the musical imagery) could be used more systematically to avoid linguistic and other shortcomings of the self-reports. Neural measures (electrical activity and blood flow, summarized by Belfi, this volume) have been applied to musical aspects of mental imagery only rarely (e.g., Herholz et al., 2012) and not yet directly connected with the theoretical and causal explanations.

Several chapters in this book do address the temporal nature of music-related mental imagery. Gelding and colleagues (this volume) describe a series of ingenious experiments where the participants report either imagery or emotions in different time points of listening in order to find out the sequential and possibly the causal order between imagery and emotions. It turns out that the emotional responses to music are usually reported before the experiences of music-evoked visual imagery. This has implications for music and emotion research as one of the central theories has outlined visual imagery as one of the key mechanisms of music-induced emotion (Juslin & Västfjäll, 2008), but of course it could be that the observed order is related to the task and that conscious attention is required to respond to emotion or imagery.

In these types of laboratory experiments (e.g., Day & Thompson, 2019; Weir et al., 2015), this field could give attention to reproducibility and replication of the findings by always replicating the results of the key studies in another laboratory through collaborations and sometimes offering the replications as student work. If such practice is not adopted, progress might be slower as the field assumes that all results are reliable, and this might contribute to slower accumulated pace of discovery or to suggest to follow leads and ideas that turn up later to be unreliable or one-off findings.

The very fact that culture shapes mental imagery does pose additional challenges to the way they are captured. Athanasopoulos (this volume) reviews the evidence of how basic parameters such as vertical or lateral movement related to music are far from universal and vary across cultures. These examples of visual and spatial imagery come from a few pockets of cultures (Papua New Guinea, Greece, Japan, and Pakistan), and the observations are often obtained in different and incomparable ways, but even with these caveats, it is a stark reminder that music-related mental imagery is likely to have significant variation across cultures and individuals. This in turns does suggest that capturing and talking about mental imagery with participants needs to be carried out in a way that is suitable for different cultures and groups of people.

Theoretical Perspective—Mental Imagery as Embodied Cognition?

The long-standing debate about mental imagery—depictive as represented by Kosslyn et al. (2006) versus propositional as represented by Pylyshyn (2002)—is still going on and there is more evidence to expand the scope of the discussion to incorporate more elements from actions and our surroundings. With this, I refer to the embodied cognition approach, which takes the notion of linking perception with action and cognition as a foundational element (Gallagher, 2006; Wilson, 2002). In its radical form (Noë, 2004; Thomas, 2014), there is a system of meanings linked to actions and interactions that each person builds by interacting within the environment. To support this framework, there is plenty of evidence suggesting that mental imagery is really not easily separable from motor and sensorimotor actions; for instance, if gaze fixations are restricted in both visual and auditory encoding tasks, it impairs the performance in recall tasks (Johansson et al., 2012), and motor processes are heavily utilized in tasks involving mental rotation (Voyer & Jansen, 2017).

If mental imagery and the types of domains involved (auditory, visual, and kinaesthetic) in this process would be considered from an embodied perspective, what would change in the planning, execution, and interpretation of the empirical studies? It would depend on what kind of embodied framework is applied as the more radical ones such as Noë (2004) and Thomas (2014) eliminate the notion of representation altogether and focus on action and motor and sensorimotor procedures that interact with surroundings. In most definitions of embodied cognition, the mental processes are linked with "modal, non-arbitrary interactions with the world, interactions that are governed by perception and action systems" (Waller et al., 2012, p. 296). Adopting this view, separate representations for imagery are not postulated but the imagery involved in music through auditory, visual, or kinaesthetic domains is interpreted as simulations of perception and action. This could complicate things in terms of having to abandon speaking and operationalizing imagery as a separate representation. On the positive side is that there are frameworks that translate and link perceptual and motor processes with abstract and conceptual knowledge such as the emulation theory of mental representation by Grush (2004), Barsalou's simulated and grounded cognition (2008, 2009), or the nonrepresentational approach to cognition by Chemero (2011). Moreover, classic research on mental scanning (Denis & Kosslyn, 1999) and conceptual metaphors (Gibbs & Matlock, 2008) has provided examples of how these links can be captured and exploited experimentally. An embodied, non-representational line of view to imagery would also bring all modalities into focus of attention in relation to imagery in a natural way as suggested by several authors in this volume (Taruffi & Küssner, this volume; Nanay, this volume; Ofner & Stober, this volume).

Conclusions

Theoretical integration, innovation, and analytical reviews of the existing findings are needed to ask more diagnostic questions about music-related mental imagery; what processes are involved, what role do the different modalities play, and how are they initially formed and acquired if they are heavily influenced by culture? In particular, the possibility of regarding mental imagery as simulation of perception and action as suggested in the embodied cognition paradigm brings the additional benefit of incorporating all modalities to the discussions and planning of this research. However, this paradigm could also burden this research theme by complicating the issues with mutually incompatible theories of enactivism (one that emphasizes the perceiver's action within the environment stressing how sensory and motor processes are inseparable) and embodiment (a broader view where cognition depends on bodily experiences within a biological, psychological, and cultural context) do not make choosing the right theoretical framework any easier. Since the majority of mental imagery related to music has involved either auditory imagery or visual imagery, other forms of imagery (e.g., kinaesthetic and motor imagery) should get more sustained attention.

There is a clear need to develop better instruments that capture music-related visual, auditory, motor, and kinaesthetic imagery. For musical imagery, the instruments need to be able to probe the musical elements (melody, harmony, rhythm, and timbre) of the phenomena and preferably to map the gross time-course of the process. Some tools, paradigms, and instruments are already available such as the Pitch and Rhythm Imagery Tasks (Gelding et al., 2015), an adaptive auditory mental imagery test (Gelding et al., 2021), and broad survey tools such as the music-related visual imagery survey by Küssner and Eerola (2019). There is more work to be done to validate such instruments with objective indices of music-related visual imagery and visual mental imagery, and the tools should be taken to cover a wider range of modalities such as motor and kinaesthetic domains. Finally, a question of how much music-related mental imagery is influenced by expertise (musical or other), individual differences, or broad cultural and linguistic differences are questions that have only been marginally explored. Overall, a music and mental imagery research programme requires evidence from multiple approaches, from validated behavioural instruments or surveys, or psychophysiological or neural measures linked to the experimentally manipulated parts of the mechanisms and triggers. This demanding aspect of the research topic suggests that collaborations, within and between disciplines, are going to be vital before this area is able to define the topic areas in a manner that allows to settle the many remaining questions.

The real-world applications of music and mental imagery already exist. There are applications in music therapy and education that rely on imagery as part of a therapy process (e.g., Guided Imagery and Music, GIM, see Fachner, this volume, and Dukić, this volume) or as a part of rehearsal and performance process (Finch & Oakman, this volume, and also Presicce, this volume). One of the possible ways to feed the basic research and promote the insights from this area is to campaign for featuring this topic more prominently in the teaching of music theory, musicology, and music performance, since these are the areas with the widest reach for people who are likely to be the music professionals in the future.

References

Baddeley, A. D., & Andrade, J. (2000). Working memory and the vividness of imagery. *Journal of Experimental Psychology: General*, *129*(1), 126–145. https://doi.org/10.1037/0096-3445.129.1.126

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617–645. https://doi.org/10.1146/annurev.psych.59.103006.093639

Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1281–1289. https://doi.org/10.1098/rstb.2008.0319

Chemero, A. (2011). *Radical embodied cognitive science*. Cambridge: MIT Press.

Day, R. A., & Thompson, W. F. (2019). Measuring the onset of experiences of emotion and imagery in response to music. *Psychomusicology: Music, Mind, and Brain*, *29*(2–3), 75–89. https://doi.org/10.1037/pmu0000220

Denis, M., & Kosslyn, S. M. (1999). Scanning visual mental images: A window on the mind. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, *18*(4), 409–465.

Gallagher, S. (2006). *How the body shapes the mind*. London: Clarendon Press.

Gelding, R. W., Harrison, P. M. C., Silas, S., Johnson, B. W., Thompson, W. F., & Müllensiefen, D. (2021). An efficient and adaptive test of auditory mental imagery. *Psychological Research*, *85*(3), 1201–1220. https://doi.org/10.1007/s00426-020-01322-3

Gelding, R. W., Thompson, W. F., & Johnson, B. W. (2015). The pitch imagery arrow task: Effects of musical training, vividness, and mental control. *PLoS One*, *10*(3), e0121809. https://doi.org/10.1371/journal.pone.0121809

Gibbs, R. W., & Matlock, T. (2008) Metaphor, imagination, and simulation: Psycholinguistic evidence. In R. W. Gibbs (Ed.), *The Cambridge handbook of metaphor and thought* (pp. 161–176). Cambridge: Cambridge University Press.

Godøy, R. I. (2001). Imagined action, excitation, and resonance. In R. I. Godøy & H. Jørgensen (Eds.), *Musical imagery* (pp. 239–251). London: Taylor & Francis.

Godøy, R. I. (2003). Motor-mimetic music cognition. *Leonardo*, *36*(4), 317–319. https://doi.org/10.1162/002409403322258781

Grush, R. (2004). The emulation theory of representation: Motor control, imagery, and perception. *Behavioral and Brain Sciences*, *27*(3), 377–442. https://doi.org/10.1017/s0140525x04000093

Herholz, S. C., Halpern, A. R., & Zatorre, R. J. (2012). Neuronal correlates of perception, imagery, and memory for familiar tunes. *Journal of Cognitive Neuroscience*, *24*(6), 1382–1397. https://doi.org/10.1162/jocn_a_00216

Johansson, R., Holsanova, J., Dewhurst, R., & Holmqvist, K. (2012). Eye movements during scene recollection have a functional role, but they are not reinstatements of those produced during encoding. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(5), 1289–1314. https://doi.org/10.1037/a0026585

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, *31*(5), 559–575. https://doi.org/10.1017/S0140525X08005293

Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006). *The case for mental imagery*. Oxford: Oxford University Press.

Küssner, M. B., & Eerola, T. (2019). The content and functions of vivid and soothing visual imagery during music listening: Findings from a survey study. *Psychomusicology: Music, Mind, and Brain*, *29*(2–3), 90–99. https://doi.org/10.1037/pmu0000238

Noë, A. (2004). *Action in perception*. Boston: MIT Press.

Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin*, *80*(1), 1–24. https://doi.org/10.1037/h0034650

Pylyshyn, Z. W. (2002). Mental imagery: In search of a theory. *Behavioral and Brain Sciences*, *25*(2), 157–182. https://doi.org/10.1017/S0140525X02000043

Tartaglia, E. M., Bamert, L., Mast, F. W., & Herzog, M. H. (2009). Human perceptual learning by mental imagery. *Current Biology*, *19*(24), 2081–2085. https://doi.org/10.1016/j.cub.2009.10.060

Thomas, N. J. (2014). The multidimensional spectrum of imagination: Images, dreams, hallucinations, and active, imaginative perception. *Humanities*, *3*(2), 132–184. https://doi.org/10.3390/h3020132

Voyer, D., & Jansen, P. (2017). Motor expertise and performance in spatial tasks: A meta-analysis. *Human Movement Science*, *54*, 110–124. https://doi.org/10.1016/j.humov.2017.04.004

Waller, D., Schweitzer, J. R., Brunton, J. R., & Knudson, R. M. (2012). A century of imagery research: Reflections on Cheves Perky's contribution to our understanding of mental imagery. *The American Journal of Psychology*, *125*(3), 291–305. https://doi.org/10.5406/amerjpsyc.125.3.0291

Weir, G., Williamson, V. J., & Müllensiefen, D. (2015). Increased involuntary musical mental activity is not associated with more accurate voluntary musical imagery. *Psychomusicology: Music, Mind & Brain*, *25*(1), 48–57. https://doi.org/10.1037/pmu0000076

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*(4), 625–636. https://doi.org/10.3758/BF03196322

Index

Locators in *italics* refer to figures and those in **bold** to tables.