2019

Making in the moment: The dynamic cognition of musicians-in-action

Kevin J. Ryan

MAKING IN THE MOMENT:
THE DYNAMIC COGNITION OF MUSICIANS-IN-ACTION

by

Kevin J. Ryan Jr.

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Major: Philosophy

The University of Memphis

December 2019

Dedication

To Mom, Dad, Aunt Mary, and Uncle Jim:
The good parts of this project are thanks to you and I take responsibility for any remainders.

Acknowledgements

This dissertation would not have been possible without an immense amount of support. First and foremost, I need to thank my parents for everything, especially in this case for helping me to cultivate a love of music that was a foundation for this project. Many thanks to Dr. Shaun Gallagher, my dissertation supervisor, whose guidance, advice, and feedback was essential in every step of progress I've made towards my PhD and beyond. Thank you to the other members of my dissertation committee - Dr. Deborah Tollefsen, Dr. Michael Monahan, and Dr. Garry Hagberg – for all of their invaluable advice and feedback. Thank you to my teaching mentor, Dr. Remy Debes, for helping me become a more confident and competent philosophy teacher. I am likewise deeply grateful to Dr. Georg Theiner, Dr. Thomas Busch, and Dr. Tillmann Vierkant for their advice and guidance during my time as a B.A. and M.Sc. student, all of which was extremely helpful during my Ph.D. studies.

Sections of this dissertation were completed during my time as a visiting scholar at the University of Exeter, Stony Brook University, and McGill University. I am honored to have had the chance to conduct research at these locations, and thank you to Dr. Giovanna Colombetti, Dr. John Lutterbie, and Dr. Laurence Kirmayer for their respective guidance and support that made those visits possible. A fair amount of time writing this project was spent at RP Tracks, and I would like to thank the fantastic staff who not only became good friends but who kindly put up with me taking over a table for hours at a time. Also thank you to the University of Memphis librarians and library staff for helping me gather the resources needed to complete my research.

Earlier chapter drafts received essential feedback from the University of Nebraska at Omaha philosophy faculty, as well as audience members at several conferences and events, including the Cognitive Futures in the Arts and Humanities conference and the IMPROV

workshop. I am furthermore honored and thankful to the Marcus W. Orr Center for the Humanities at the University of Memphis for the Freeburg Dissertation Fellowship that allowed me to focus on dissertation research in Spring 2019, and the MOCH fellows who provided both guidance and feedback on my work.

Finally, it would be an understatement to say that I am indebted to Connie Diffie and Cathy Wilhelm for everything they have done for me and the philosophy department throughout my time at Memphis. I cannot thank them enough. Similarly, thank you to all of my UofM colleagues in philosophy and beyond, including but not limited to the philosophy faculty and graduate students, audiology faculty and graduate students, and Institute for Intelligent Systems faculty and graduate students. Last, but certainly not least, thank you to all of my other colleagues, family members, and friends around the world. You have all been an amazing source of support, love, and inspiration. You have made me a better scholar, thinker, and person, and I can never truly express how much that means besides saying thank from the bottom of my heart.

Abstract

Ryan, Jr., Kevin J. PhD. The University of Memphis. December, 2019. Making in the Moment: The Dynamic Cognition of Musicians-in-Action. Major Professor: Dr. Shaun Gallagher, PhD.

Watching highly-skilled experts in the midst of improvised performance can be a source of mystification and wonder. Understanding this mystery in more detail, especially for music, is a major motivation for my dissertation. Moreover, while there are multiple avenues through which one could explore improvisation, I will primarily utilize the tools of embodied cognitive science to help better understand it in general and bebop jazz improvisation in particular. I furthermore consider possible ways to define improvisation as an essential precondition of my project.

In what follows, I will not defend one type of embodied approach over all alternatives. Instead, I will consider improvisation in light of three different strands of research: shared intentions, ecological psychology (especially in regards to a theory of affordances), and predictive processing. The focus on these strands, taken both as individual research programs and a single unit of analysis, closely mirrors the essential core commitment of embodiment by providing a dynamic account that spans brain, body, and world. It likewise does so in ways that reject any neat partition of inputs, cognition, and output.

For shared intentions, the main issue I will explore concerns how to best account for the dynamic moment-to-moment engagement of musicians with each other, especially in light of improvisation as an intentional activity, and covering the differences from novices to experts in bebop performance. For a theory of affordances, the focus will be on the interplay between a skilled agent and a structured environment as an essential part of musical perception and action. Finally, for predictive processing, a picture of the brain as a predictive, anticipatory engine takes center stage to explain how musicians can respond to the extremely fast time constraints that are

part of musical performance. I will also consider how novelty can be accounted for on predictive processing accounts. Through considerations of these different areas, the impacts of embodied approaches will be further clarified and help us to better model, understand, and appreciate the cognition of musicians-in-action.

Table of Contents

Introduction

Watching highly-skilled experts in the midst of improvised performance can be a source of mystification and wonder. From the perspective of a novice, one may further ask how does a group get together - in front of an audience, no less - to make interesting, engaging, detailed, and sometimes even beautiful results? As Paul Berliner has noted for jazz improvisation (1994):

> For those outside the jazz community who discover improvisers as mature artists through their recordings, these issues [e.g. how does an artist build and maintain a distinct improvisational style?] remain mysterious. For prospective musicians who wish to follow in the footsteps of their idols, however, unraveling the mystery is essential. (p. 2).

Understanding this mystery in more detail, especially for music, is a major motivation for my dissertation. Moreover, while there are multiple avenues through which one could explore improvisation, I will primarily utilize the tools of *embodied cognitive science* to help better understand it in general and bebop jazz improvisation in particular.

Exploring the link between music, improvisation, and cognition has several important theoretical and practical upshots. First, doing so helps address the issue of how best to understand the conceptual and cognitive grounds for various skills needed to perform music, regardless of skill level. Second, doing so helps address pedagogical concerns about how to develop fluency and expertise in music-making (de Bruin, 2015; de Bruin & Harris, 2017). Third, doing so helps establish which skills are task specific for music and which generalize to different contexts (Beaty, 2015; Pressing, 1998).

Addressing the extent to which cognitive capabilities underlying musical improvisation are domain-specific or domain-general has additional ramifications for the use of jazz

improvisation as a metaphor in non-musical contexts. The metaphorical use of jazz has seen particularly strong application in recent work on management theory, for example (Kamoche, e Cunha, & da Cunha, 2003). On one hand, management theorists have highlighted the need for agility in responding to unforeseen situations, just as a jazz improviser fluidly and actively responds to a rapidly changing sonic environment. On the other hand, there have been critiques of this uptake as playing into dubious assumptions about the music, which in turn potentially serves to support domination by management in the workplace and ignores important parts of jazz's sociohistorical roots (Hatch & Weick, 1998). Furthermore, such problematic assumptions sometimes rest on an incorrect view of jazz improvisation as a purely individual and infinitely flexible practice. As we shall see below, and in line with some of these critical concerns, jazz improvisation is as much a collective, collaborative, and transformative process as it is about learning how to flexibly fit into existent modes of performance and practice.

Music and Cognitive Science

Understanding cognition is a unifying task for cognitive science. There is, however, no unified sense of the exact role that music should play vis-à-vis the study of cognition. To some, music is unfit for inclusion as a topic for cognitive scientists to study in its own right, and better seen as a mere offshoot of other faculties, most notably a side-effect of linguistic processing (Pinker, 1997). For others, music offers a distinct and essential part of the human cognitive story, especially in relation to grounding the early development of language as well as being an important part of human evolution and sociality (Cross, 2001; Mithen, 2005). Christopher Longuet-Higgins, who's considered the first to have officially termed the phrase "cognitive science," was one early proponent of studying music cognition, both in its own rights and to

better understand cognition in general. (Longuet-Higgins 1979; 1994; also see Laske, 1977; Pearce & Rohrmeier, 2012).

In addition to pondering the place of music within cognitive science, and regardless of where it ultimately sits in the field, a distinctive psychology of music has developed in recent decades. One of the first and still influential treatises in this literature is Fred Lerdahl and Ray Jackendoff's *A Generative Theory of Tonal Music* (1983). While other influential accounts have been developed over the years (e.g. Krumhansl, 1990; Deutsch, 2013; Narmour, 1990; Meyer, 1956), I shall briefly introduce Lerdahl and Jackendoff's Generative Theory of Tonal Music (GTTM) here to give a general sense of the kinds of explanations that have thus far developed in the psychology of music.

The main target of the GTTM is capturing how a trained listener understands and categorizes a piece of Western tonal music. Drawing heavily from linguistics work on generative grammar (Chomsky, 1965), Lerdahl and Jackendoff suggest that people must have innate, rule-like structures that allow them to parse musical surface material. Well-formedness rules are used to track notes, chords, and rhythms of a piece, along with several other hierarchical features. At the same time, since there are occasions where multiple interpretations are "correct" according to well-formedness rules, they cannot be the entirety of the story. They are thereby joined with preference rules that establish how to adjudicate conflicting well-formedness rules as needed. Well-formedness and preference rules, in turn, explain why and how a listener picks between different but equally acceptable alternatives to reach a final understanding of the song.

Taken together, both kinds of rules purportedly capture how an implicit knowledge of musical grammar[1] underlies our ability to interpret a piece of music. Just like the native speaker

[1] Musical grammar traces the syntactically acceptable combinations of notes in tonal music, akin to the syntax of a natural language.

who is able to create grammatically acceptable sentences as well as evaluate sentences, the listener fluent in Western art music will have a sense of "good" and "bad" musical creations. While the GTTM is not meant to explain everything,[2] it paints a clear picture for how essential aspects of music cognition function along the lines of the assumptions built into generative linguistics.

We also find rules-based assumptions in regards to some attempts to model improvised musical performance. Philp Johnson-Laird's (2002) model of improvisation draws on the idea that agents apply several different kinds of rules or algorithms during performance. Testing of this model has resulted in computer agents successfully creating improvised solos, albeit at the level of a novice performer. While not necessarily the same sort of rules as found in GTTM, these algorithms follow a rule-based approach for the creation of musical material, which in turn can be related to a particular understanding of cognition as a computational process. In Johnson-Laird's words, regarding the study of creativity, "whether or not creativity is computable is an open question. A sensible strategy is accordingly to assume that it is computable until one is forced to abandon this hypothesis. In contrast, a *theory* of creativity ought to be framed in a computable way" (2002, p. 418).

Disembodied Music Cognition

I suggest that the GTTM and Johnson-Laird's model are exemplars of *disembodied music cognition* (cf. Linson & Clarke, 2017; Kersten, 2017). While Lerdahl and Jackendoff are attuned to the importance of music in relation to the sorts of bodies we have, along with our constantly developing socio-cultural and material contexts – for instance, the innate mechanisms in their universal musical grammar adapt to local musical contexts, and Charles Nussbaum (2007) has

[2] Timbre and other non-hierarchical features are noted and commented on briefly but left to the wayside, since the GTTM is primarily concerned with hierarchical features of music.

provided additional evidence that the GTTM may be psychologically implemented in the human motor system instead of an abstract logical space - their focus on cognition as abstract rule-processing in brains implies that all the relevant cognitive work occurs inside the individual, in principle locked away from the world and the specifics of the bodies that we have, just like a disconnected brain-in-a-vat.

The specific features that mark a theory as disembodied remain a live debate. Following a fairly general consensus, I suggest that disembodied theories are at least tied into two main theoretical commitments. First, they operate under the assumption that the immense amount of information in the world, internal and external to the agent, requires a heavy dose of inferential work in order to become meaningful. In the words of William James on infant experience, the world is here taken to be a "blooming, buzzing confusion" before we build it into a coherent whole in our minds. This assumption is tied into "poverty of the stimulus" arguments (Chomsky, 1980), wherein input from the world is taken as too poor to provide the basis of skills such as language acquisition without something more, and usually innate, at play inside the agent.

Second, disembodied theorists hold that the organism-based processing performed on this immense yet poor information operates over a rather strict distinction between sensory inputs, core cognition, and motor outputs (i.e. action). Following Susan Hurley (2008), we can call this the "sandwich model" assumption. The sandwich model is similarly related to modular approaches to cognition (Fodor, 1983), where all that really matters for cognitive science happens at a particular level of abstraction away from the body, or the latter at best merely implements the relevant cognitive processes without doing much of the most important work. Both of these commitments are part and parcel of *computationalist* and *cognitivist* approaches to

classical AI, to which we now turn in more detail.[3] According to proponents of cognitivism, which is explicitly part of the framework for Lerdahl and Jackendoff's work (also see Raffman, 1993, chap. 2), the neatness of the breakdown between sensation, cognition, and action entails that, while all three are relevant for a given cognitive process, the *real* core of cognition can be treated without any specific references to the body, apart from the fact that there is some body that receives sensory information and facilitates motor output. There is likewise a historical privileging of certain processes, such as decision making, planning, and abstract reasoning, as the most important targets for cognitive science. In its strongest form, cognitivism entails a view of the mind as primarily geared towards thinking about abstract truths and relying on rational processes of decision making, in contrast to being an evolved organ with a primary purpose of aiding an organism's survival in a rapidly changing environment (Fodor, 2008).

The cognitivist picture also implies that perception starts by detecting lower-level features of a sensory stimulus before moving to more complex features in a step-wise fashion. Seeing a house, in this case, would begin by seeing lines and edges, then seeing more concrete particulars of the object, and finally seeing a house – perhaps even my house in contrast to your house - as such. One immediate issue with this claim, as Eric Clarke has noted, pertains to a conflict with phenomenological experience. When listening to music, Clarke notes that (2005):

> Direct experience suggests that this [cognitivist or information-processing picture] is wrong: if you hear a burst of music from someone's radio, for instance, it is more likely that you will be able to say what style of music it is (opera, hip-hop, Country and Western) than to identify specific pitch intervals, or its key, meter, and instrumentation. In other words, people seem to be aware of supposedly

[3] In what follows, the use of computational or cognitivism without any modifiers will be used interchangeably with disembodied.

> ‘high-level’ features much more directly and immediately than the lower-level features that a standard information-processing account suggests they need to process first. (p. 15-16).

While phenomenological evidence such as this example offer important data for consideration in a theory of music cognition, a defender of the cognitivist approach would likely reply that the conscious experience of finding a genre before pitch intervals may not reflect the unconscious, subpersonal flow of information involved in cognitive processing. It is similarly possible for the brain to rapidly process lower-level features of a song and reach mid-level and higher-level features without consciousness being impacted one way or the other along the way, since the speed of information processing may occur at a time below the threshold of consciousness.

Pursuing this computational approach further, a classical philosophical formulation can be found in the work of Jerry Fodor (1980). For Fodor, our best psychological theories operate on two sets of assumptions: a representational theory of mind and a computational theory of mind. According to the representational theory of mind, all mental states operate over representations. These representations are distinguished by either their content or the nature of the relation between subject and content. For instance, the difference between “Miles believes that it is raining” and “Miles desires that it is raining” comes from the difference in the mental state belief vs. desire, since the content is otherwise the same. In contrast, “Miles believes that it is raining” and “Miles believes that it is sunny” bear the same relation between subject and content, since both are instances of belief, yet they are different mental states since their content is different.

For a computational theory of mind, mental processes only have access to syntactic properties of internal mental representations. Fodor refers to this claim as the ‘formality

condition'. According to the formality condition, mental state operations proceed without reference to semantics; notions such as truth, reference, and meaning are therefore unacceptable as part of a psychological theory. As a result, all that matters for the work of cognition takes place internal to the organism. Hence Fodor developed an additional claim about "methodological solipsism" for cognitive psychology, according to which psychological theorizing should advance with a strong internalist assumption. Likewise, it entails that the relevant sorts of psychological theorizing about the mind can and should take place without appealing to the world outside of the organism, and perhaps even further limited further to just the organism's brain.

How does this general Fodorian argument about representation and computation relate to disembodied music cognition? As seen above, and as noted by Diana Raffman (1993, p. 26-27), the cognitivist account of understanding music entails that music perception operates under increasingly abstract processing of stimuli from physical input, through cognition, to output (usually with output being an understanding of the piece). The computationalist account not only fits well into this picture, but it adds additional restrictions as to the kinds of information available throughout musical processing.

Embodied Music Cognition

Embodied cognitive science, which stands opposed in various ways to the two core assumptions of traditional cognitivist and computational approaches listed above, has gone by a plethora of acronyms since its rise in the late 1980s and early 1990s. In what follows, I shall interchangeably refer to embodied cognitive science using the term "4E paradigms" or just "4E" approaches, all of which take cognition to be a fundamentally *embodied*, *embedded*, *enactive*, and *extended* process.

All 4E supporters agree that understanding cognition requires taking seriously the dynamic interaction between brain, body, and world. The duration of these dynamic connections includes: the milliseconds of neural firings, the seconds of immediate conscious awareness, the years of individual ontogenetic growth, and the eons of phylogenetic development. They also concern how events in these different time-scales interact with each other (see Gallagher, 2017, chapter 1). Such a dynamic interaction among and within different timescales necessarily breaks down any clean partition between input, cognition, and output. It likewise puts pressure on cognition being distinct from, and nested between, perception and action. Furthering research in embodied music cognition requires continued development and application of the various theoretical and empirical tools used to explore exactly how this rich interplay can realize various cognitive phenomena, scaling from simple to complex tasks (Elliot & Silverman, 2015; Schiavio, 2015; Van der Schyff, Schiavio, Walton, Velardo, & Chemero, 2018; Leman, 2007; 2016).

To further ground the nature of an embodied approach to music cognition,[4] Lawrence Shapiro's general taxonomy of cognitive science is an important additional tool to use (Shapiro, 2010, p. 5-6). It is constructed as a series of Yes/No questions, beginning with whether the two approaches – i.e. 4E cognitive science and disembodied cognitive science - have the same subject matter. Since our focus is on music cognition, the answer to this point is clearly yes. We are then faced with whether or not 4E and disembodied approaches offer *competing* explanations. As noted by Shapiro, if these approaches are direct competitors, we should adopt the one that provides the best explanatory framework. If not, then either approach may be adopted by practicing scientists in the field and pursued accordingly.

A proponent of embodied music cognition may be content that their approach is seen as a

[4] Different definitions of embodied music cognition have carved up the relation along different lines, just as different theorists prioritize different constellations of E's in their approaches.

useful parallel option to disembodied approaches; in addition, both avenues may have independent motivations for their continued development without a markedly different explanatory goodness in adopting one kind of account. Yet even if we take the next step and instead say that, yes, embodied music approaches are in direct competition with disembodied music cognition approaches and, furthermore, 4E is preferable, that alone does not address the swath of differences among embodiment theorists regarding, among other things, the best kinds of psychological explanations and the structure of cognitive ontology in basic and more advanced forms of cognition.

The 4E paradigms were historically presented along a unified front. Part of this unification was a pragmatic choice by those working within the field. As they have continued to develop, however, it has become clear that not everyone within 4E cognitive science agrees on many issues outside of the basic idea that understanding the brain, body, and world interconnection is fundamental to understanding cognition (Kiverstein & Clark, 2009). This failure to agree among 4E theorists is not necessarily the mark of a failed research project but instead a sign of a robust, lively research field that continues to develop. Yet it also requires us to consider several live issues within embodied cognition in more detail, both to better understand music through its lens and to offer important case studies in music that feedback to research in embodied cognition writ large.

In what follows, I will not defend one particular type of embodied approach over all alternatives. Instead, among the multiple possibilities at hand, I will consider improvisation in light of three different strands of research that can be situated within 4E cognition: shared intentions, ecological psychology (especially in regards to a theory of affordances), and predictive processing. The focus on each of these strands, taken as individual research programs

and taken as an overall unit of analysis, closely mirrors the essential core commitment of embodiment by providing a dynamic account that spans brain, body, and world in ways that reject the neat cognitivist partition of inputs, cognition, and output.

For shared intentions, the main issue I will explore concerns how to best account for the dynamic moment-to-moment engagement of musicians with each other, especially in light of improvisation as an intentional activity, and covering across the differences from novices to experts in bebop performance. For a theory of affordances, the focus will be on the interplay between a skilled agent and a structured environment as an essential part of musical perception and action. Finally, for predictive processing, a picture of the brain as a predictive, anticipatory engine takes center stage to explain how musicians can respond to the extremely fast time constraints that are part of musical performance. Through considerations of these different areas, the impacts of embodied approaches will be further clarified and help us to better model, understand, and appreciate the cognition of musicians-in-action.

Chapters Overview

In Chapter One, I begin by considering head-on what is improvisation? Doing so is essential for making sense of the relevant literature on music improvisation and improvisation in general. After surveying a variety of possible definitions, I provide a two-tiered understanding of the phenomenon. These tiers track a general explanation of improvisation on one level and a specific explanation of jazz improvisation on the other. The remainder of the chapter consists of grappling with various issues for my approach, such as the relation between improvisation and planning, the role of success conditions in jazz improvisation, and how to relate spontaneity and randomness to improvised activity.

Chapter Two begins with considerations about the sorts of intentions that are part-and-

parcel of bebop ensemble performance. I therefore consider both individual and shared intentions. In doing so, I likewise explore several accounts of shared agency to find one that provides the necessary conceptual tools to understand bebop performance across all levels, from those just starting to play to the pioneers and giants in the genre. I will suggest that some of the most developed accounts of joint and shared agency – including the work of Margaret Gilbert and Michael Bratman - provide a good explanation for part, but not all, of the different levels of performance. I conclude by defending the dynamic theory of (shared) intentions as originally introduced by Elisabeth Pacherie, and which I take to be the best available option for my target phenomenon, against several enactivist critiques that demand an even more minimal cognitive explanation than the one provided in this theory.

In Chapter Three, I consider how the tools of ecological psychology can be used to analyze Coleman Hawkins' 1939 performance of *Body and Soul*. The main model I utilize is Mark Risjord's ecological model of improvisation. After introducing it, I turn to critically analyze and further develop its three central components of attunement, affordances, and meta-cognition. More specifically, attunement will be taken in conjunction with the nature of ecological information; affordances are approached within metaphysical debates about how best to understand them; and meta-cognition is focused on the claim that "the brain is a resonance organ," both in regards to whether or not it can (1) make sense as more than a metaphor and (2) offer an alternative to contrasting accounts of the brain as an organ that represents the world.

In Chapter Four, I turn to consider anticipation in music. In particular, the theory of predictive processing will be the main issue under analysis, along with the core issue of how the demand to decrease uncertainty that sits at the heart of predictive processing can, nevertheless, leave a space for novelty. Two different frameworks that extend the main tenants of predictive

processing – the Active Inference Framework and the Cultural Affordances Framework - will also be considered here, since they offer two distinct ways of bringing the predictive processing account into contact with other research areas.

Each of chapters two, three, and four, while occasionally making reference to other areas of the dissertation, are primarily considered in (semi-)isolation. In the conclusion, I bring them together in a brief yet direct dialogue to ask about how we can model the cognitive processes underlying jazz improvisation. In short, I suggest that frameworks of predictive processing provide a key part of the cognitive story regarding the brain and cultural development; ecological psychology helps better explicate the environmental structures along with some of the internal structures that make jazz improvisation possible; and shared intentions, along with different theories of shared agency, make sense of the moment-to-moment unfolding of musical improvisation in an ensemble. I finish by considering additional upshots of these ramifications and posit forthcoming questions for continued research.

Chapter One: What is Musical Improvisation?

Music can make us move and groove. We tap our toes, bop our heads, and snap our fingers to an infectious beat. We interact through complex forms of engagement as well, humming tunes, writing lyrics, singing karaoke, and playing air guitar. Sometimes we move to music with great reserve, meekly approaching a potential partner at a school dance; other times we move with reckless abandon, moshing at a heavy metal concert or stage diving at a punk show. Likewise, some movements are highly patterned and synchronized, such as line dancing at a country concert, while others are not, such as contact improvisation in a dance studio.

We continuously find across these cases that music is tightly interwoven with bodies in the world. As Tia DeNora (2000) has noted, however:

> In the realm of common sense, the body is paradoxical. At once self-evident and mysterious, biologically 'given' yet modifiable, the body is characterized through contradiction. Exploring these contradictions helps to open many deeper questions concerning the relationship between bodies and the material-cultural settings of their existence. (p. 75).

Following in her footsteps, I join with others to continue taking seriously the task of understanding the relationship between the body, the material socio-cultural world of music, and their multifaceted realization in human life (Small, 1998; Clarke, 2005; Krueger, 2009; 2011; 2014; Moran, 2014; Leman, 2007, 2016; Nussbaum, 2007; de Souza, 2017; Cox, 2016; Borgo, 2005; Overy & Molnar-Szakacs, 2009).

A complete understanding of this nexus between body, music, and life requires considerations of moment-to-moment musical engagements. As seen in the examples above, some of these are highly structured while others take place in a free-form manner. There are

important differences across these cases, no doubt. Yet we can state with certainty that musical activities of all kinds, for performers and listeners alike, inherently rely on a variety of *improvised* activities.

For our immediate purposes, we can understand improvisation in light of concepts such as spontaneous, on-the-fly, and/or non-planned activity. Sometimes improvisation is part-and-parcel of the aesthetic goals of performance, as in the case of a jam band or a small bebop quartet. Other times it's not. Yet even within exemplars of ostensibly non-improvised musical forms – e.g., a modern-day symphony orchestra playing works of the standard Western classical music repertoire in a concert hall - there remains an important role for improvisation (Gould & Keaton, 2000).

Orchestra musicians, like all other musicians, are necessarily responsive to elements of different performance spaces. Any given performance will include various flourishes and changes across time. Such responses are not all drawn up in advance, even across hours of copious practice and with the use of a score. These different patterns of responsiveness - which could take place among musicians, between musicians and the audience, or between audience members themselves - emerge spontaneously in the moment and help create subtle nuances and differences night after night. Furthermore, insofar as they arise spontaneously, they are improvised in an at least one important sense of the term.

In contrast to current views on Western classical music, one musical genre commonly associated with improvisation is jazz. While jazz history cannot solely be marked by improvisation – doing so would exclude large segments of early jazz, for one, and big bands where composers and band leaders alike left little space for improvisation during performance, for another – the two are often thought of as thoroughly intertwined. This connection is most

clearly seen in subgenres such as bebop and its offshoots, including cool bop and hard bop, as well as the case of free jazz and related referent-free improvisational practices.

Before delving more into the concept of improvisation, however, it will help to say a few words on the common form of bebop performance. The style first developed in the early 1940s with artists such as trumpet players Dizzy Gillespie and Miles Davis, pianist Thelonious Monk, saxophone players Charlie Parker and John Coltrane, and drummers Kenny Clarke and Max Roach. While not immediately accepted by everyone, it has since become one of the most important developments in the history of the genre, fundamentally shifting our understanding of jazz as a whole. In the words of Scott DeVeaux, for instance, bebop "is both the source of the present—'that great revolution in jazz which made all subsequent jazz modernisms possible'—and the prism through which we absorb the past. To understand jazz, one must understand bebop" (1997, p. 3).

Bebop musicians take a *standard* as the core of their performances, which could be composed explicitly for that purpose, drawn from popular tunes of the time (commonly from Broadway or musical theatre) or passed down in the jazz cannon from previous eras. Often developed around a 32-bar, 16-bar, or 12-bar blues format, the standard begins with the band playing what is called a "head". During the head, a melody is stated (or implied), usually by being played once or twice with the entire band. Next, the song moves into the solo section. Solos are harmonically constrained by the chord changes introduced in the head, somewhat open rhythmically, and extremely open melodically. After everyone has taken a solo or two, the band goes to the outro where the head is restated again and the song ends.

It is moreover important to note that, even when a bebop musician is not taking a solo, they are nevertheless constantly improvising throughout the performance. This continual

improvisation is a marked difference from earlier big band and swing era jazz, which were epitomized by bandleaders such as Count Basie and Duke Ellington, and where most of the performance was centered around prewritten compositions with at most a small opening of prespecified spaces for a performer to improvise.

Preliminary Attempts at Defining Improvisation

At the start of this chapter, I suggested a preliminary set of related concepts for improvisation: on-the-fly, spontaneous, and non-planned activity. While useful placeholders, these concepts alone do little to help parse out the distinct conceptual space around improvisation. They also don't help us better understand the sorts of cognitive processes underlying improvised actions. Furthermore, figuring out what counts as improvisation – indeed, even partially establishing what improvised activity is compared to non-improvised activities - remains an extremely difficult task. Many common definitions tend to focus more on what improvisation is *not* rather than what it *is* (cf. Berliner 1994, p. 1-3). Improvisation *is not* planning; improvisation *is not* composition; improvisation *is not* a direct reproduction of a product from memory. While these definitions may say something important about improvisation, they don't offer us nearly enough to get a sense of what improvisation actually *is*.

Positive definitions of improvisation often center on concepts such as spontaneity, freedom, novelty, originality, and uniqueness (Alperson, 2010). As noted by Andrew Goldman (2016), in addition to theoretical discussions, this set of concepts has played a strong role in various operational definitions of improvisation used in empirical literature. Another empirical focus on the nature of improvisation takes it to be a play with preconceived motifs or structures, which Philip Johnson-Laird has termed the "motif theory" (2002). To be more specific with these kinds of positive definitions, several recent experiments about musical improvisation have

defined the phenomenon as follows: “Spontaneous musical performance, whether through singing or playing an instrument, can be defined as the immediate, on-line improvisation of novel melodic, harmonic, and rhythmic musical elements within a relevant musical context” (Limb & Braun, 2008, p. 1); “In terms of cognitive processes, improvisation can be defined as the spontaneous generation, selection, and execution of novel auditory-motor sequences” (Berkowitz & Ansari 2008, p. 535); “[Improvisation] involves freely generated choices” and “must be adapted to ongoing performance, and monitored through auditory and somatosensory feedback, as well as to an overall aesthetic goal” (Bengtsson, Csíkszentmihályi, & Ullén 2007, p. 831).

While these and similar, earlier definitions have played an important role in the history of discourse around jazz, especially in romantic notions that link improvisation with genius and spontaneous creativity (Gioia, 1988), an immediate problem one faces is the extreme difficulty of establishing when it is or is not acceptable to attribute them to a given performance (cf. Goldman, 2016). Even if we assume an objective measure for something like novelty or creativity, any ascription of what counts as a novel performance will, at best, be made *post hoc*.

Likewise, different metrics, especially when we consider the development of taste over time, may entail that a single performance could be both novel and not-novel or spontaneous and not-spontaneous *at the same time*. It may be possible to appeal to some kind of pluralist approach among what counts as novelty or spontaneity, and it is possible to locate and mark these concepts in purely observer independent terms if we, e.g., define novelty in relation to previous performances by the same performer or in relation to how similar or different it is to common performance practices at the time. Doing so, unfortunately, just pushes the issue back without explaining what marks something as improvised, since it can still be ostensibly improvised and

not improvised at the same time. In order to address these concerns, a better approach seems to be that one ought to stay as far away from these particular concepts as possible, at least when it comes to defining improvisation, and even if we may ultimately use them to evaluate the quality of different improvised performances.

An alternative option to the above positive and negative definitions, following a classic philosophical treatment of improvisation provided by Gilbert Ryle (1976), starts by casting an extremely wide net around various types of improvised activities in order to better understand the phenomenon. In his words, improvisation is that element which for an improvisor (1976):

> can be the good qualities of their wits and characters that come out in their impromptu jokes, their swift repartees, the fresh conclusions they draw, their shots at the goal or the ways they cope with traffic-emergencies. It is part of intelligence to seize new opportunities and to face new hazards; to be, in short, 'not a tram, but a bus'. (p. 69).

The tram/bus metaphor is particularly instructive in this case and I shall embellish on it here and again below. For starters, a tram is able to travel either forward or backwards along its tracks, thus leaving a certain openness available to it, but its path remains heavily circumscribed from the outset. The restricted nature of the situation leaves tram drivers unable to respond to many aspects of the environment except in preset ways, mainly involving the choice between one of two directions and either increasing or decreasing speed. As a result, while the driver can speed up, slow down, and “turn around” by going in reverse, any serious debris on the track can leave a tram stranded for an indefinitely extended period of time.

In contrast, a bus may find a new route when faced with road blockage. While it may not able to travel everywhere - steep, narrow mountain passes and off-roading in a field may be

unavailable to even the most adventurous bus drivers, after all – it is able to go to a much wider set of areas and has more maneuverability than the tram. At the same time, the bus driver must be more responsive to various features of the environment. Debris in the path may not be as much of a problem for the bus as for the tram, yet the increase in freedom also increases responsibility at the same time.

Moving away from this metaphor for the moment, and although specific examples of music are lacking in Ryle's account (cf. Alperson, 2017), it's important to note that his approach entails we can say something similar for any and every sort of improvised activity. Jokes, jazz, and conversation are all treated on the same level in this kind of account. Keeping them so seems important insofar as they are all cases of improvisation, and a safe assumption is that there at least has to be something making them all tokens of the same type. However, since my interest is ultimately on the relevant cognitive processing required in specific kinds of improvised activity, there is a worry that being too broad with the content of a general account may be problematic when it comes time to answer specific what and how questions about improvisation.

This concern also mirrors a longstanding criticism of behaviorists from the likes of functionalists. Accordingly, if our focus is solely on inputs and outputs that black box the processes occurring in-between – i.e. if we only focus on course-grained similarities between activities that actually require a variety of different skills and cognitive processes - we may miss extremely important differences in the cognition underlying distinct kinds of improvised activity, in this case. A large breadth has its pros, to be sure, but it has some cons as well.

For a more concrete example, consider the relationship between music and language. These are commonly referenced as being similar in multiple ways, such as in the recruitment of neural regions and, in turn, related cognitive capacities. For instance, there is strong evidence for

overlapping cognitive processes drawn from fMRI studies that have found activation in similar neural areas, such as Broca's area and Wernicke's area, which are all active during musical performance and language comprehension (Fadiga, Craighero, & D'Ausilio, 2009). Likewise, studies on the dynamics of improvisational solos have considered a role for the non-verbal dynamic patterns between a soloist and their accompaniment that have similarities to the non-verbal feedback cues in a conversation (Moran, Hadley, Bader, & Keller, 2015).

There are important differences between the two processes, however, such as the fact that conversation between too many simultaneous speakers breaks down in a way that music played among an equally large number does not. Music production also often occurs when multiple musicians play at the same time. Indeed, a group playing a song actively requires multiple people performing simultaneously for the majority or entirety of the performance. A conversation where everyone is talking at once, in contrast, is doomed to failure.

An additional study by Christopher Kello, Simone Dalla Bella, Butovens Médé and Ramesh Balasubramaniam (2017) reaches similar findings by considering hierarchical structures of music, speech, and animal vocalizations (specifically whale and bird song), which gives additional evidence for similarities and differences between language and music. Hierarchical structures were focused on here since they are a common feature across these different categories. The authors in turn found a "natural taxonomy" among these structures that allowed them to categorize certain important similarities and differences. For instance, social interaction seems to increase the clustering and structure of sounds produced in the timescale of seconds, regardless of whether the social interaction was in speech, music, or whale song. In addition, the temporal hierarchical structuring of jazz performance was more closely related to speech than it was to either Western classical or pop music, adding additional credence to the theory that jazz

improvisation is like a conversation, perhaps even more so than other styles of music.

The results also found even steeper clustering in the case of music recording than in the case of speech, which the authors note could not directly be explained by the social interaction thesis. I suggest that one possible read of this point, in line with the previous several paragraphs, may instead stem from the nature of the social interaction occurring in music. In conversation, only one person is often speaking, while in music all musicians are contributing at the same time. These differences most likely result in different dynamic profiles across the two contexts, even though some similarities remain. Combined with genre differences, such as those among classical and pop as noted in the study, may further explain the higher amounts of clustering in recorded music over recorded conversation.

This brief consideration of music and language does not establish any deeply significant differences or similarities between the two, largely for limitations of space. I am furthermore not taking any direct stand on the language-music link here. What I hope to have instead provided is a brief but important set of considerations for pushing our analysis deeper to the particular features of different improvisational activities. To return to the tram/bus metaphor, we know that the world isn't solely populated by trams and busses. Moreover, there are different sorts of skills required, and unique possibilities made available, with driving a tram, bus, car, motorcycle, or scooter, among a plethora of other options for transportation.

Another approach to defining improvisation, suggested by George Lewis and Benjamin Piekut in their introduction to *The Oxford Handbook of Critical Improvisation Studies*, is to reject the task of providing a preset definition in the first place. While not opposed to defining improvisation in principle, Lewis and Piekut are concerned that any *a priori* definition wouldn't allow fruitful development for ongoing, interdisciplinary research into the various nature and

inner workings of improvisational practices. In their words, after surveying a wide swath of different improvised activities, "What emerges from this extended, yet necessarily incomplete, discussion of issues is the futility of drawing boundaries around the critical study of improvisation" (2016, p. 14). Taken in another way, we can understand this option as one where we forego definitions and let the examples do the work. Instead of fretting over any particular definition, these authors suggest that we should spend our time understanding how different kinds of improvisation operate on their own terms.[1] Any relevant definitions of improvisation will only come about after doing so, and may only be useful for limited cases of particular study instead of as broad categorization tools.

I agree with Lewis and Piekut that there are good reasons for not being overly consumed with a strict definition of improvisation, especially if it is taken in terms of necessary and sufficient conditions. Doing so may, as they note, foreclose unforeseen avenues of exploration within and between disciplines. This foreclosure could therefore be a loss in the short and long terms. In addition, existing positive definitions of improvisation sometimes run into problems by considering certain features as necessary for improvisation, as was discussed above. These problems include a concern for *post hoc* ascriptions of certain features to an improvisation and a potential problem of delimiting the phenomenon without taking into account culturally mediated aspects of different improvised practices. Furthermore, a jazz solo, driving in traffic, and running a last-minute committee meeting may all be improvised activities, but there is no common collection of necessary features that these activities share.

[1] It is important that we situate this claim in the context of Lewis and Piekut's immediate project, which is to offer a wide-ranging introduction to a variety of explorations of improvised activities that are both disciplinary and interdisciplinary in scope. However, we can take their stance as a general point for an account that wishes to say something about improvisation in general. I take this second option as my guide and continue to do so for the remainder of this section.

Nevertheless, and with this concern notwithstanding, a failure to offer any definition, or at least a schematic definition, first threatens to breakdown our understanding of the upshots we ought to keep specific to jazz-related cognition and the upshots that can be expanded to cover other forms of improvisational activity. Second, following the examples-first approach also makes it difficult to explain why we should or should not consider these various examples under the umbrella of improvisation. Judging what counts based on intuition and a gut feeling that this activity is, in fact, improvised – as in I'll just know it when I see it - may give us a potential way out for this second concern, but it would fail to address the first.

This final point about grouping improvised activities together brings back a major hurdle to defining improvisation that has returned again and again in my discussion so far. This issue concerns the scope of analysis, specifically if a working definition should attempt to capture *all* types of improvisation under one concept or not. While there are certain *prima facie* connections between different improvisational activities, I am not convinced that these connections entail some strong cognitive link between them. On a similar note, regardless of how broad or narrow the ultimate story we tell about a cognitive practice is, as well as whether or not improvisational skills and practices turn out to have more domain-general or domain-specific capabilities, I suggest that we need to provide an interrelated, multi-level definition to explain any given improvisational practice. The dual questions addressed in this multi-level analysis concerns why a particular phenomenon is improvisational in the first place and how specific types of improvisational activities function in particular contexts.

The first level, which concerns what makes something improvisational in broad strokes, will provide a definition writ-large that addresses why activities such as improv comedy, jazz, a spontaneous conversation, a witty reply, a detour, etc. are all types of improvisation. It will thus

provide a litmus test of sorts for a mark of improvisation.[2] As a general definition, it will abstract away from some of the fine-grained details across concrete cases of improvised activity.

The second level, which concerns the context of a specific improvised activity, will offer an explanation that more clearly lays out how improvisation works within that particular practice. While there will presumably be overlap between different improvisational activities, it's an open empirical question regarding the extent of this overlap and how it matters for psychological and cognitive explanations. The question of how similar or different these sorts of explanations should be is thus an open one that is a task for continued discussion and research.

General Definition of Improvisation

The *general definition of improvisation* demarcates the phenomenon in its broadest strokes. I suggest that this definition should at least be implicit in any study of improvisation, although it is not necessary to be explicit in all contexts; it's of most interest in philosophical or theoretical considerations. As such, this definition will be one that brings together the myriad of cases referenced in the excerpt from Ryle above. Or, to run more closely to the tram/bus metaphor, this general definition should focus on what makes improvisation a particular kind of vehicle, without saying something about how Hondas and Toyotas are different brands of cars or busses. As a result, I suggest a general definition of improvisation as *the navigation of an intentional activity within a particular environment and without strict adherence to a fully preset plan*.[3] This definition is clearly broad enough to capture multiple types of improvisational activities.

[2] One could raise the question of whether or not this kind of definition may ignore important cultural differences regarding what marks something as improvised, especially insofar as deeming something to be improvisation carries normative commitments along with psychological differences. As such, the definition of improvisation I provide here should be read in a more value neutral, and perhaps even conditional, terms. In other words, the definition can be prefaced with "if an activity is to be considered improvised, then it will…". Whether or not the term improvisation should be applied may also be a matter of other sociocultural and historical considerations.

[3] In a similar manner, Vijay Iyer, reflecting on the relationship between improvisation and temporality, suggests that improvisation can be understood as "the real-time interaction with the structure of one's environment" (2004, p. 164).

Moreover, it acknowledges that, while parts of improvisation draw from memories and preexistent cognitive structures, just as navigating an environment requires drawing on knowledge of and/or previous experience in that environment, the product created through improvising cannot be completely predetermined in advance. Likewise, this definition highlights and entrenches the fact that many of our actions, from the most basic to complex, include some improvisational components.

Such a definition faces three immediate issues that need to be addressed. First, especially since it does not utilize concepts around novelty or spontaneity, what is the relationship between improvisation and randomness? Second, what exactly is meant by "strict adherence to a fully preset plan?" Third, if our general definition of improvisation implies most everything is improvisation of some kind, what role is left for non-improvised activity? In other words, is this definition ultimately too broad in its scope?

Turning to the first issue, I grant that randomness can be a *part* of an improvised activity, but I nevertheless hold that randomness, in itself, should *not* be *equated* with improvisation. One example for why can be seen in the difference between an actor falling by accident and the same actor performing a prat fall. A well-executed prat fall looks identical to what would happen if a person had tripped and fallen by mistake. By its nature, however, a prat fall will not be random nor accidental. It is an intentional activity that only happens to look accidental from an external perspective. The accidental fall, in contrast, will be random as well as non-intentional, and thus not a candidate for being an improvised action (at least not until it receives uptake from either the actor or others around them to continue moving the scene forward). For an analogous jazz case of a random accident becoming improvised, consider the phenomenon of playing a random, "wrong" note multiple times. The case of the repeated "mistake" serves to take the randomness

out of the original note in a way that marks the overall activity as non-preplanned but still intentionally guided.

We can further consider going to see a play together. This play happens to include a scene where two people walk down a street in the midst of a heated argument. No fall was planned in the original script, and no fall has been scheduled to occur during this particular exchange. Suddenly, however, the first interlocutor trips and falls, adding a comedic moment to an otherwise tense scene. The fall also happens to befit the overall bumbling nature of the character, who is left flustered but carries on with the conversation without missing a beat.

Now let's turn to a bit of intuition pumping about this case. If, after the show ends, we find out that the actor had mistakenly tripped and fallen mid-sentence, we would most likely think of the fall simply as a mistake, even if it were an ultimately fortuitous one for the overall quality of the show. We may also positively judge the actor recovering from the mistake, yet it seems to me that we would still not view the fall itself as an act of improvisation. It was, after all, something that merely happened at the time, instead of something that was done.

In contrast, if the actor had decided during the scene that they wanted to add a moment of levity to the performance and had taken a prat fall, the situation would be different in some important way. We therefore would, or should, think of that fall as an improvised part of their performance from the start, not a mistake around which people had to improvise to get the scene back on track. While in both of these cases we may positively or negatively evaluate the fall – its acceptability, its usefulness, how successfully the actor and their co-actors responded to it, etc. – only one of these cases would be an improvised fall, while the other would simply be a mistake.

This example opens up several additional points about the nature of improvisation in relation to randomness. First, while the case of the mistaken fall may not in itself be a case of

improvisation, the responses from the actors would necessarily involve some improvisation to continue the show. As John Lutterbie (2011) has noted in regards to theatrical improvisation:

> To improvise successfully requires inhibiting the existing cognitive processes and distractions that interfere with the task at hand and being open to letting the conceptual blends run. When this frame of mind is achieved, a chain of associations spirals out as image begets image. It is not a free flow of material, however, because the actor is continually making choices that move the improvisation in difficult directions, based primarily on the attraction immanent in what is happening at that particular moment. (p. 178).

The inhibition process referenced here occurs in both fall cases from above. Yet the way in which the falling actor decided to act is an important distinction between the two.

Second, the nature of the prat fall raises questions about improvisation and temporality (Iyer, 2004; Lewis & Piekut, 2016). The majority of improvisational activity under analysis in bebop jazz concerns situations that are created in, and responsive to, immediate features in the environment. As Vijay Iyer (2004) has noted in terms of research on the auditory processing of music, an improvising musician needs to respond in a window of around $1/25^{th}$ of a second. Yet improvisational activities may also occur outside the context of the immediate moment. Thus the "immediate environment" in the general definition is not limited to the moment of immediate conscious awareness, but can spread over larger periods of time and, in some situations, across a rather large geographical space as well.

Third, the prat fall highlights the issue of when does something become improvised when it wasn't before. In following the prat fall vs mistaken fall cases, for instance, it would be a mistake to scrutinize the scene for the exact moment where improvisation was taking place. Did

it start when the actor decided to take the fall? Did it start when they actually took the fall? Did it start when the other actors began engaging with the actor who had fallen down? A similar issue can be raised in forms of improvisation that take place over longer periods of time, such as in the case of hiring deliberations among a committee or distributed processes of social activism, which largely move at time scales far outside the millisecond and second levels of musical and theatrical performance. The solution to these questions, as a result, is to hold that it is a mistake to look for a magical moment that changes a non-improvised activity into an improvised one.

Before returning to this point below in regards to the second major issue for my definition vis-à-vis preset plans, there is an additional problem to be faced here concerning the significant one-way control inherent with the idea of improvisation as an "intentional activity." Put succinctly, since some important cases of improvising only operate when a performer opens themselves up to outside forces in the environment, the worry is that my general definition, with a demand for intentional guidance, forecloses the possibility of clearly improvised activities from being considered so in the first place.

As a first response, I suggest that it is possible for someone to direct themselves to be open to the world. In doing so, they relinquish control and allow the music, scene, or other elements in the environment to take the lead and drive the performance into unforeseen and unintended locations. At the same time, while the performer gives up most or even all local control over what happens, they do so without relinquishing overall control over the situation. As a result, these sorts of activities can both fit into the general definition and do so without jettisoning the requirement for guided, intentional control.

This first response, unfortunately, is not yet fully satisfactory. How unsatisfactory it is will become more noticeable if we consider the case of musical absorption. Absorption in

performance occurs when musicians lose themselves in the music. It can also occur without them deciding to, planning to, or actively choosing to relinquish any control over their actions. There are many different types of musical absorption, but the strongest comes from deep absorption (see Høffding, 2019). In these cases, should we understand the deeply absorbed musician as guiding their improvised activity through the use of intentions? Is musical performance in deep absorption taking place without any intentions guiding the performance choices being made in the moment?

In order to satisfactorily address this new iteration of the problem, we need to briefly shift our focus from what counts as intentional activity to what it means to navigate an environment. An analogous case to the deeply absorbed musician is the absent-minded driver, who may drive for miles before realizing, sometimes with a shock, that they have no idea how they got to a specific spot on the highway and five or ten minutes have passed without notice. Like the absent-minded driver, the fully absorbed musician still remains aware in essential ways about the music and other activities going on in the performance space. Even if they are not aware of how they are partaking in it over a period of time, they are continually taking part in a skilled and guided navigation of the music. Perhaps this sort of navigation is not available to their conscious reflection, and it may never be consciously available whenever deep absorption occurs, but that does not mean that no navigation is taking place. Furthermore, deeply absorbed navigation, while remaining guided, is also taking place without strict adherence to a preset plan, thus fulfilling the latter half of the general definition.

We can now turn to the second major concern for the general definition concerning how we should understand that improvisational activity does not allow for “strict adherence to a fully preset plan”. Returning to the prat fall example in light of this specific criterion, one could turn

my accidental fall-prat fall argument on its head. This inversion happens if we agree that accidental trip has less connection to a preset plan than the prat fall. If so, then the accidental fall would be *more* of a candidate for improvisation than the preplanned prat fall, which goes completely against my stance above.

The quick way to diffuse this concern is by adding the qualification that the actions relevant for improvisation must be undertaken by an agent. The random fall may be less connected to a preset plan than a prat fall, of course, but it fails to qualify as a relevant action in the first place. As a result, we can exclude it in contrast to the prat fall. I will return again to consider a similar sort of issue in Chapter 2 on the relation between intentions and action.

For a more complete response, we find that the apparently strict connection between action and preset plan actually comes in degrees. The two poles of this continuum are completely random action on one end and completely preplanned action on the other. If an action is at either pole, it will not be improvisational. The complete disconnect at the random pole is encapsulated in the accidental fall, taken in itself and without reference to the larger context. The other pole is exemplified in a waiter who never deviates from their script when greeting a new table, an actor who similarly refuses to deviate from the script, or a musician who can only play what is presented by a score. The number of actions that will actually take place at either pole is rare and, and especially at the fully preset plan end, may not exist in practice.

As a result, we can now turn to the third and final main issue for my definition concerning what non-improvisational activity is, if every action is on the spectrum of improvisation. In other words, this is a worry that the general definition is just too broad and problematically entails that all activity is improvised by definition, in contrast to finding general activities that have some sort of improvisational components.

In response to this concern, I suggest that we focus on the difference between the claim "all activity involves improvisational elements" and the claim "all activity is improvisational in kind." For an analogous case, biologists make distinctions in kind among different species, even though the same mechanism of change through natural selection has driven development throughout all different cases. In this analogy, improvisation in the general definition is akin to the mechanism of natural selection, while the difference between species is offloaded onto the specific definition (for, e.g., bebop performance).

In the case of improvisation in general, a similar source material can furthermore have more or less improvised versions. Take the case of *My Favorite Things* as one such example. There are clearly important differences between a musician performing *My Favorite Things* in its original Broadway form and as a jazz standard developed by John Coltrane. However, even in the case of staying closely attached to the Broadway standard, there remains important improvisational elements at play about specific choices made in any given performance, even if they are not in the same degree as the Coltrane version of the song.

While the above may have addressed some of the worries of improvisation bloat, there is still an additional concern from the opposite direction that adherence to any rule, strictly held or not, may collapse my definition in the opposite direction. The case of free jazz, for instance, still involves the use of a variety of rules, even if their content will be negotiated in the moment of performance itself. Moreover, if these rules are sufficiently strict enough, then one may hold onto the idea that the activity will similarly be fully captured by these kinds of rules.

There are at least two responses to this point. First, it is important to distinguish between rules and plans. Rules and constraints are a necessary component of activity, but they are not the same as plans for those activities. We will consider the nature of how plans are related to

improvisation in more detail in Chapter 2, so I will not say more about them here. Second, insofar as improvisation is a scalar notion between two poles, the rules at play and the way of relating to them will have an impact on where an activity falls along the spectrum, but it will not be enough to push the activity off it entirely. In other words, the rules will shape the way in which an improvisation unfolds, but it will not make the activity non-improvised.

Explanatory Definition of Jazz Improvisation

According to Richard Ashley (2009), the cognitive processes involved in musical improvisation are best understood within the context of several kinds of constraints. He notes three particularly pertinent limitations on body, time, and knowledge storage, which are in turn explained as physical, temporal, and cognitive constraints during performance. Recent empirical work in the area is also overlapped with concepts such as creativity and "flow" state (Landau & Limb, 2017), although whether we want to build creativity into the definition of improvisation, in contrast to the idea that it has a positive aesthetic goal, remains an open issue.

In order to address these concerns in the case of bebop performance, I suggest that we understand jazz improvisation as the *navigation of a jazz performance environment with relevant environmental and socio-cultural musical affordances towards the goal of a successful improvised musical performance*. Even on a quick reading, a concern of circularity is immediately present with this particular definition. In order to get around that concern, we can understand "improvised" in the definition as shorthand for the general definition. As a result, the full definition of jazz improvisation is the *navigation of a jazz performance environment with relevant environmental and socio-cultural musical affordances towards the goal of a successful, intentional musical performance without strict adherence to a fully preset plan*.

As with my general definition of improvisation, I take this claim as a start instead of the

end of my analysis. For instance, the nature of affordances and other features of the environment will be the detailed focus of Chapter 3, so I will not so much about that part of the definition at the moment. With that caveat in place, there are two additional questions that must be addressed before moving forward to the remainder of my analysis. First, and to return to a discussion that began above, what is meant by navigation in this particular context? Second, what denotes a successful performance and why include success as part of the definition of jazz improvisation, instead of packing that into the additional consideration of what makes a good jazz performance?

The notion of navigation is relevant here for two reasons. First, it highlights important ways in which the improvisational process is not a series of random happenings, as discussed for the general definition above. Moreover, it foregrounds the fact that a musical context is one that is already positioned in a socio-historical context with a variety of available options within a sonic environment to choose from, along with various sets of know-how needed for possible ways to operate.

In navigating a dense city landscape, for example, there are a variety of different ways to move from point A to point B and reach one's destination. Some of these paths are straightforward while others are winding and time-consuming. There will also be paths that only appear to a person who knows the city well and to which a newcomer may be unfamiliar with, just as there is an important difference between what an expert musician can do during performance and what a novice is only beginning to comprehend.

A jazz improviser, regardless of being a novice or an expert, will play in a musical landscape where they must react to various features in a continuous manner. Based on a combination of skill, practice, and personal idiosyncrasies, a musician will respond to the demands of real-time development of the music, just as a navigator in any transportation must be

able to navigate through various sorts of environments in relation to the tasks at hand. In a similar fashion, it takes an extended period of time from learning how to operate a vehicle to becoming a professional bus-driver. It likewise takes years of practice from learning an instrument to reach the standing of a professional musician.

Second, the concept of navigation highlights an important social element of bebop and jazz improvisation. Even the most basic tasks of navigating are not something that individuals do by themselves. We first learn to navigate through others. We likewise use the accumulation of various knowledge, techniques, and technologies from others to move about in our environments; indeed, at the same time as we shape these environments with new technologies and social practices, they simultaneously shape us.[4]

In addition to these points of consideration, the notion of navigation has been commonly used to discuss jazz improvisation. For instance, it was employed by David Sudnow in his classic text *Ways of the Hand* (2001). In this text, Sudnow offers a phenomenological reflection on his development as a jazz piano player, with a focus on how this learning process occurred. In particular, he suggests that learning to improvise is about embodying certain possibilities for moving about the sonic space, such as using particular scales and chords, plus learning how those skills could be used to navigate the piano both as a solo player and within an ensemble.

An important addition to the notion of navigation can be made in connection to Tim Ingold's (2000, Chap. 13) distinction among different ways to map an environment. Two main options are available in this practice. There are processes of *cartography* that are tied into abstractly capturing the contours of a location from a generalized view from

[4] This bidirectional process between organism and environment, known as niche construction (Laland, Odling-Smee, & Feldman, 2000; Laland, Matthews, & Feldman, 2016) in biological literature, is considered in more detail below.

everywhere/nowhere. This type of map-making is about plotting locations in space and charting out the possible relations or paths between them. In addition, there are also processes of *wayfinding* focused on navigating an environment that has a particular history for a particular person. When wayfinding, a person is not thinking about their environment as a generalized and abstract geometrical plane but, rather, as a series of places within particular regions. As a result, the process of wayfinding is integrally connected to forms of story-telling, which includes directly relating to the sedimented history in the area.

To a certain extent, we can see the two different processes of mapping related to George Lewis's (1996) distinction between different logics of improvisation: eurological and afrological. The eurological approach, with its focus on technical rules and regulations, holds close to the generalized processes of cartography and map-making. The afrological approach, with its focus on self-realization and society, holds closer to wayfinding. At the same time, just as there is no intrinsically better or worse way to map one's environment, my claim about this relationship is not that one approach is in itself superior to the other. However, there are essential differences between the two, and additional avenues for understanding how these processes differ is an important part of a better coming to terms with various assumptions about the world, ourselves, others, and their interrelation.

The mapping distinctions can be further modified by adding a third category of wandering. All of these processes – cartography, wayfinding, and wandering - are modes a jazz improviser could uptake when navigating their performance. Jazz pedagogy and its focus on helping musicians develop specific skills may make it possible to move from the rigidness of cartography, where one is seeking out abstracted paths that already exist, to wayfinding, where one must understand and engage with the musical material and their distinct position in the

world. It may ultimately even show how wandering of certain kinds may help in both of these situations.

Wandering crosscuts cartography and wayfinding, and is integrally tied into both processes. It allows for an individual or group to find new options to move about their environment, altering the sorts of sedimented patterns that can happen in abstract map-making and concrete wayfinding. Wandering, in other words, can shake up the map legend and alter the boundaries of a map over time. Whether or not wandering is a good thing will rest on a variety of additional considerations. In the context of music-making, these include whether or not we are asking about the wandering process from the position of the performer during an improvisation or from the perspective of the audience who is listening to that same improvisation. It also includes whether the musician is attempting to epitomize the heights of a preestablished genre or critically develop inroads in a new, untested areas.

Wayfinding furthermore entails the possibility of restructuring a landscape where one is navigating at the same time as taking part in the process of navigation, in a similar manner to stigmergic path formation. This process of path formation is a self-organized one where paths become created over repeated use by people over time; however, the planning is not necessarily created from a centralized decision, but will instead emerge more spontaneously and, through repeated uses, become more sedimented over time (Theiner, Allen, & Goldstone, 2010). Cartography likewise opens a bidirectional relationship between the environment and how we navigate it, since it is both beholden to landmarks and features that already exist while, simultaneously, having a map reorients how we understand the relations between those elements in the world. The main difference between the two concerns how we develop our understanding of the environment in which we are navigating, as noted above.

With navigation more clearly explained, we can turn to the last bit of the particular definition of jazz improvisation concerning a successful performance. This concern can be broken into two additional points. First, what counts as success? Second, if success is a constitutive feature of jazz improvisation, how do we account for flawed solos that have other notable aesthetic traits or bad improvisations that remain improvised and within jazz despite their aesthetic and technical flaws?

In terms of the first point, there is no single criterion for success that can be distilled from philosophical reflection alone. Some of what would be considered a success for a bebop performance depends on the time of the analysis. For instance, there are important differences understanding success between when bebop first developed into its own subgenre, when it had reached maturation by the late 1940s and 1950s, looking back at it as a crystalized moment in musical history, and considering people who still play bebop today. There are also different considerations for success, such as emotional depth, technical mastery, or responsiveness to other bandmembers, that go above and beyond the issue of "wrong" notes being played. An important part of applying the definition of jazz improvisation will thus require weighing different sociocultural expectations that vary from audience to audience and performer to performer. Doing so should not be done from the philosophical armchair. Going this route also gives leeway for talking about distinctions within different sorts of jazz improvisational activities and different jazz subgenres, as well as capturing the development of distinct sorts of jazz improvisation based on different developments within the genre.

Rejecting any *a priori* definition of success does not account for the second specific issue of how to capture unsuccessful solos or bad performances. To do so, I take success in the definition, not as a necessary condition that must be reached, but as first and foremost the *aim* of

jazz improvisation. For the case of a novice, striving towards a successful performance is part of the learning process. For an expert, a more stringent standard of what counts as success may be applied. In either case, and no matter what those standards ultimately are, we find that the notion of success is integral to how we situate and evaluate different performances.

There is an important difference among improvisational practices within jazz, moreover. For instance, an audience will have a widely different set of expectations if they went to see a big band, bebop, or free jazz performance. Developing skills as a performer goes hand-in-hand with an audience developing aesthetic tastes and listening habits to appreciate different performance practices. Exploring whether or not these sorts of improvisational activities are similar or different mirrors the concerns about the relation between improvisation in general and musical improvisation in particular. Settling that issue is outside the scope of my current project. Nevertheless, to keep from accidentally shifting between these different performance spaces, I will conclude this chapter with a brief taxonomy of improvisation as originally developed by Lee Brown, David Goldblatt, and Theodore Gracyk (2018) and relate it back to my specific explanation of jazz improvisation.

A Further Taxonomy of Jazz Improvisation

Brown et al. (2018, p. 194-200) introduce five categories for capturing the different kinds of improvisation that occur in jazz writ-large. An important feature underlying all five categories concerns whether or not the musicians intend to perform, or at least begin a performance, with a particular tune in mind, on one hand, or if they do not draw inspiration from any specified source material, on the other. It is furthermore important to note at the outset that these categories are non-orthogonal and it's possible to find overlap between them within the same performance.

With these caveats in mind, we can now turn to the categories themselves. The first two – (1) expressive individuation of a recognizable work and (2) improvisation sanctioned within a score-directed performance – both closely follow work-specific properties during the process of crafting an improvisation. They do so either by respecting the thin work elements for a song in category (1) (e.g. while there are specific melodies associated with a tune, some songs don't specify the requisite tempo or instrumentation and, therefore, may be played at different speeds with a variety of different instruments yet remain instances of the same song) or in (2) by improvising when, and only when, the score calls for it.

The third category – (3) Improvisation based on a composed or recognized work – the musicians intend to begin with a recognizable work but ultimately deviate in large ways from it. In this category, the song as composed is an important ingredient, but it is neither maintained during the performance nor do the musicians intend to respect the identity features of the work as was the case in categories (1) and (2). One example, as Brown et al. mention, is Coleman Hawkins' 1939 version of *Body and Soul*, which I shall explore in more detail in Chapter 3.

The fourth and fifth categories – (4) Spontaneous musical development of non-work ingredients and (5) free improvisation – account for performances where there are no references made to work-relevant identity conditions. For instance, in the case of (4), musicians may begin with a 12-bar blues without attempting to emulate a specific work when doing so. In (5), no intentions are made to follow any referent during performance, although local constraints may arise and fall at various points.

Both parts of my two-tiered definition of general improvisation and jazz improvisation above offer something in relation to these five categories. At the general definition level, these categories can be seen as different points along the continuum between the poles of

improvisation. At the specific level, these five categories offer a more concrete sense of what success conditions may be at play for a given performance. Sometimes those conditions can only be applied *post hoc*. However, unlike a strong ideal of novelty and the concerns that came with it for a positive definition of improvisation, my analysis does not rest on them for distinguishing between a real and a fake improvisation; it only matters for distinguishing a good vs. bad one.

Chapter Conclusion

In this chapter, I considered issues with some of the dominant ways for conceptualizing improvisation currently available in the literature, which includes negative definitions, relying on concepts such as novelty or spontaneity, and letting the examples do all the work. I then proposed that any analysis of improvisation will rely on two different definitions, the first being a general definition and the second a specific case of the target phenomenon at hand. After providing a definition for both the general case and the case of jazz improvisation, the remainder of the chapter was spent clarifying some potential concerns with those levels, as well as briefly relating them to an existent taxonomy of jazz improvisation in the literature to show how they can all work together.

In many ways, the remainder of this dissertation can be read as doing the work of both fleshing out additional parts of the specific level of jazz improvisation as well as seeing how this particular area has ramifications for our understanding of cognition, both in general and in light of bebop jazz performance. Furthermore, as was apparent in the Brown et al. taxonomy, there is an important connection drawn between improvisation and the intentions of a performer. I will turn to those intentions in the next chapter.

Chapter Two: Improvisation, Knowledge, and Intentions

Improvisation is far from a mindless, random activity. One way of understanding what guides it, including the on-the-spot decisions musical improvisation entails, grants an essential role to a musician's intentions during performance. To address this approach, in addition to considering the nature of intentions in general, we must understand how musicians relate themselves to the intentions of others. In order for a group to successfully improvise together, it is important that musicians monitor a wide variety of individual and collective features of performance, which includes their own cognitive states, their immediate phenomenological experience, and the cognitive and emotional states of others (Keller, 2008; also see Pacherie, 2012 and Tollefsen, 2014 for collective cases in general).

Noting such differences does not yet establish the best way to understand either intentions in performance or the relevant monitoring processes. Are the intentions solely a case of individual intentions? Are they some kind of collective or shared intentions where the whole goes beyond the sum of its parts? If it's the latter case, are the shared intentions underlying ensemble bebop performance reliant on more cognitively robust or cognitively simple accounts of joint action?

Among these alternatives, Garry Hagberg (2016; 2017) has argued that an important part of a jazz performance cannot be reduced to the heads or individual minds of each musician. He thus defends an irreducible conception of group jazz performance. In his words, (2017):

> given that we are talking about not only improvisation, but indeed *group* improvisation, any model that would carry with it the implication that as an intentional action we could foresee the entire enactment of the intention, or to put it another way, envision the full action in question as we tend to think we can or

could do of any individual autonomous action…would be a nonstarter. (p. 302, emphasis in original).

We shall return in more detail to this point below as we explore the various ways in which these non-individual intentions may be conceptually understood and psychologically realized.

Besides better understanding the nature of shared intentions and joint action, a focus on intentions during jazz performance has two philosophical benefits. First, in relation to the philosophy of action, exploring intentions in this context offers an important case study concerning how prior intentions and intentions-in-action are related. While the latter kind of intention has received general philosophical attention over the last several decades (Searle, 1983; Bratman, 1987; McDowell, 2008), there are still pressing questions to be addressed. In addition to questions about individual intentions-in-action, the focus on ensemble bebop performance means that we can further consider the possibility of establishing collective intentions-in-action during group performance, over-and-above individual intentions-in-action.

Second, in the sphere of philosophy of art and aesthetics, there are important debates about the role of intentions in relation to constituting musical works, some of which have focused on the question of whether or not there is such a thing as a jazz work (Brown, Goldblatt, & Gracyk, 2018, chap 7). Moreover, in relation to the categories of improvisation explored in the previous chapter, Lee Brown, David Goldblatt, and Theodore Gracyk have argued that "There are different *kinds* of improvisation, distinguished by performers' intentions about their playing" (2018, p. 200). As a result, having a clearer sense of what intentions are present in jazz performance will have something to add to these debates.

Turning first and briefly from the question of intentions as such to joint action and shared intentions, there are two main sorts of accounts available for the task of explaining the latter

processes. The first, which I shall refer to as "maximalist" accounts, are exemplified in the work of theorists such as Margaret Gilbert (1989; 2014), Michael Bratman (2014), John Searle (2014), and Raimo Tuomela (2007). While each of these accounts is distinct in important ways, they're linked insofar as they consider joint action to take place in a conceptually and cognitively robust manner. Part of this robustness is grounded in the fact that their respective starting places are neurotypical adult humans fluidly acting in the world, often performing complex tasks such as painting a house together, taking a walk together, making a committee decision, or singing a duet together.

In addition, maximalist accounts are robust since they require the agents in the joint action to have detailed amounts of information available about their co-actors. They furthermore require that the individuals in question have the ability to understand mental states, track the mental states of others, and conceptualize parts of the actions for themselves and others. As a result, besides starting with cases between human adults, these accounts have poor track records in dealing with joint actions involving young children or nonhuman animals.

In contrast, "minimalist" accounts are less cognitively and conceptually demanding. As John Michael has noted in an overview of joint action and music, "Such approaches are often called 'minimalist,' since they start out from broad (i.e. minimally demanding) definitions of joint action" (2017, p. 160). These sorts of accounts have proven useful for capturing phenomena such as non-human animal interactions and adult-infant interactions. It is furthermore possible that these sorts of cognitively less-demanding accounts may explain the nature of the cognitive processes underlying more apparently difficult joint actions, such as performing together in a

bebop ensemble.[1] Whether the best account is ultimately maximalist or minimalist, another important point in either case is that we don't praise a performer who constantly stumbles by accident onto interesting melodic lines. We praise them insofar as a large part of their performance was under some level of intentional control. It is also possible that the overall shape of intentions may change in absorbed vs. non-absorbed contexts – the musician may fail to explain the intentions behind their choices when absorbed, for instance, and they may be attending to different features of the performance than when the performance requires salvaging in notably non-absorbed contexts - but the point remains that there is always some kind of attending and intending going on in order to keep the piece in control and moving forward. As we shall consider in more detail below, there are features of intentions that are not necessarily salient in one's phenomenological awareness, yet still present in their actions.

Appealing to intentions also serves a role in distinguishing among different kinds of movements and actions that are part of improvised performance. Whether or not I intentionally step on my dance partners foot while dancing the tango matters, for example. This distinguishing role of intentions, in turn, charts important differences for groups as well as individuals, especially if we consider the difference between a collective and a mere group of unrelated individuals acting in apparently similar ways. To offer a concrete example of intentions in group cases, following Searle (1990), we can compare two similar scenarios of people gathered in a park. First, consider a crowd of individuals running for shelter as it starts to suddenly rain. In this case, we can assume that everyone is intending to run to the shelter, and the form of each person's intention is something like "I intend to reach this shelter to escape this rain." Second,

[1] A similar point has been raised in Interaction Theory approaches to social cognition regarding the role of primary intersubjectivity as both an ontogenetically early development in infants and the main mode in which people engage with others throughout their lives (Gallagher, 2001).

consider a dance company who throws performance in the same park, starting in a spread-out formation and sprinting towards the shelter at the same moment. In both cases, we can postulate that the same exact patterns of movement are at play. As a result, to an external observer, everything about what unfolds before them is identical. Nevertheless, while each individual performer in the second case may also hold a variety of individual intentions during the performance (e.g. "I intend to do my part of the performance"), something would be missing if we limit our description to just these kinds of intentions. Perhaps what is missing is something like an intention with the form "We intend to perform this dance as a company" or "I intend that we complete this dance as choreographed"

I suggest that what is missing in the second case is the distinctive group component captured in the latter two statements of intentions. A similar result holds in musical ensembles. While there may be individuals within such groups who receive more credit, such as the bandleader or soloists, it is impossible for one person to complete the entire activity on their own. Likewise, while an individual could use technology to perform all pieces of a string quartet and stich them together into a recording that sounds just like a four people playing at the same time, there is something different happening in a series of solo performance and a quartet of musicians simultaneously performing together.

Noting that difference between intentional activity in both of these contexts, however, does not yet motivate the need for some sort of robustly shared or collective intentions, instead of a mere concatenation of purely individual intentions. As a result, I first turn to explore the nature of individual intentions in some detail, with a particular focus on the role of intentions cashed out in the rapid, real-time speed of actions like taking a prat fall or taking part in a dance performance. These kinds of intentions have gone by several names in the literature, including

intentions-in-action (Searle, 1983), present-intentions (Bratman, 1987), and proximal-intentions (Pacherie, 2008).

I then turn towards specifically shared intentions and joint action. In doing so, I return to further motivate Hagberg's position on shared intentions as necessary for group jazz performance by considering how robust the cognitive mechanisms are needed to explain the joint action underlying ensemble jazz performance. I do so by drawing from conflicting maximalist versions of Gilbert and Bratman, as well as a more minimalist version in the shared dynamical intentions model as developed by Elisabeth Pacherie and Deborah Tollefsen. The minimalist account will then be juxtaposed to an enactivist account from Andrea Schiavio and Simon Høffding. While maximalist and minimalist accounts ultimately offer something important for understanding different parts of ensemble bebop performance, I suggest that the minimalist version from Pacherie and Tollefsen does the best job of providing the necessary cognitive foundation of shared intentions across the variety of bebop performance contexts.

Individual Intentions

What are intentions? They are, for starters, some sort of mental phenomenon. More specifically they are often taken as a kind of mental state. Some philosophers, such as Donald Davidson, have suggested that intentions and intentional activity are best understood as the combination of beliefs and a pro-attitude, such as desire (Davidson, 1963). Others, including Bratman, suggest that intentions should be held as a distinct kind of mental state with distinctive normative import for practical reasoning above and beyond what can be grounded in a mere concatenation of beliefs and pro-attitudes (Bratman, 1987).

To ground the distinction more concretely, consider the following three claims: (1) "I believe there is a beer in the fridge"; (2) "I desire to drink a beer"; and (3) "I intend to drink a

beer." Each of these three claims has similar content concerning beer, however, they operate in importantly different ways. In regards to (1), I may be right or wrong about my belief concerning the beer in the fridge, but that fact alone won't drive me to action either way. After all, I have plenty of beliefs about the content of the fridge, even at times when I am neither hungry nor thirsty, and these beliefs don't in themselves have any impact on my plans or actions in relation to immediate or future fridge-related behavior. For (2), the desire to get a beer would provide good reasons and motivations to seek out and drink one. Yet the desire itself doesn't include any information about where to find a beer. Thus, if all I had were the desire to drink a beer and no beliefs about where to find one, there would be no reason for me to go to the fridge in an attempt to quench this particular thirst.

The difference between the Davidsonian and Bratmanian approaches to intention can be seen in how we relate the statements about (1) and (2) to the final statement (3). Based on an early Davidsonian reading, intentions are merely complexes of beliefs and desires (or similar pro-attitudes). An agent who has (1) and (2) will therefore have (3) and, accordingly, go get a beer from the fridge. Bratman, in contrast, has argued that there are important reasons that intentions cannot be defined merely through such belief-desire complexes. A failure for me to get up from the couch and walk to the fridge for the beer where I had (1) and (2) may or may not lead you to question whether I really had the relevant belief or desire, but there isn't anything intrinsic to either the belief, desire, or their combination through which I would necessarily go to the fridge. In contrast, if I claim that I have the intention to get a beer and remained on the couch indefinitely, you may actively judge my failure to get up. This ability to pass judgment is a key to the relationship between intentions and means-end, practical reasoning, according to Bratman.

There are likewise times when one may desire contradictory things – e.g. I desire to go to

the gym and I desire to take a nap, both held simultaneously – without raising any questions of consistency among these desires. In contrast, holding a set of contradictory intentions – I intend to go to the gym and I intend to take a nap, both held simultaneously – would necessarily imply either a failure of practical reasoning or a situation where you only truly intend to do one or the other, since doing both at the same time would be impossible. For these reasons, I suggest that we treat intentions as irreducible to mere belief-desire complexes, although there are ways in which beliefs, desires, and other pro-attitudes are relevant for understanding intentions.

Thus far I have been considering intentions more or less as if there were only a single kind of intention. There is good evidence, both conceptual and empirical, however, that suggests there are in fact several different kinds of intentions. A common way of addressing this distinction is through dual-theory accounts. For these theories, one type of intention is aimed at things that will happen in the future. This type, in turn, has been referred to by several different names: future-intentions, prior-intentions, or distal-intentions. The other type is aimed at things that are happening at the moment, and have been referred to as intentions-in-action, present-intentions, or proximal-intentions (McDowell, 2011; Pacherie, 2008; Bratman, 1987). Proximal-intentions are indexical and context-dependent, since they are directly tied into the implementation of particular actions within particular environments.

While these dual-theory approaches are a step in the right direction, there is still more to intentions that needs to be explored to understand the cognition involved in bebop performance. Following Elisabeth Pacherie's (2008) dynamic theory of intentions, also referred to as the DPM account (Mylopoulos & Pacherie, 2018), I suggest that there are three different types of intentions: Distal Intentions (D-intentions), Proximal Intentions (P-Intentions), and Motor Intentions (M-intentions). The addition of motor intentions fleshes out the specific

implementation of P-intentions, which is important for understanding intentions during improvisation.

It may first seem that the only relevant distinction at play between these three types of intentions would be temporal. On such a reading, D-intentions are concerned with actions that have not yet taken place; P-intentions are concerned with either implementing D-intentions when the time is right or guiding actions that began without any preexistent D-intention; and M-intentions are concerned with guiding motor implementations of P-intentions to realize specific features of an action. For example, looking back at the prat fall case from the previous chapter, a decision to take a prat fall tonight, if made before the moment of performance, would be a D-intention, and the execution of that prat fall would entail P-intentions and M-intentions. The mistaken trip, in contrast, would not entail any D-intentions, even if responding to it would require the use of P-intentions and M-intentions to keep the scene successfully moving forward.

We should, following Pacherie, push against this simplistic temporal reading, however. The key mark between the different levels is not whether or not the intention is about things to come vs. things that are currently ongoing. Relevant differences also exist for the formation of the intention, how one moves between different sorts of intentions, and the global dynamics that weave them all together. There are additionally good reasons to assume that if D-intentions are present at the outset of an action, they will remain throughout its execution, especially in more complex actions.

D-intentions originate in deliberations on what ends ought to be pursued, paired with practical reasoning towards attaining those ends. While context is a relevant consideration for all intentions, D-intentions are not closely tied to the immediate context in which they are formed, especially when they are aimed at situations that can only be completed in the future. Their

content is partly conceptual and descriptive. When an action is put into play, and assuming that a D-intention sparked the start of that action, they maintain rational constraints, controlling and shaping the relevant forms of practical reasons while the action unfolds, including means-end coherency with a global goal in focus, along with forms of acceptable conduct for the agent.

P-intentions are concerned with starting an action now. Sometimes they may be derived from D-intentions, while other times they are not. Either way, P-intentions are at least partially perceptual and indexical. They operate at the intersection between the abstract conceptual level of D-intentions, the current situation, and knowledge about relevant motor responses that can and should be brought to bear to navigate the action(s) taking place within the current situation. P-intentions have temporal bounds that reflect the nature of the action in question, both insofar as they take place in the moment and since they must track the distinct "tempo" of the action, whether it's quick as a second or as long as a few minutes.

M-intentions operate underneath the level of conscious awareness and reflect an additional layer of guidance and control over one's actions.[2] The content of M-intentions are motor representations, which have three specific characteristics: their representational format is stored in a manner that is well-suited to motor actions; they contain some sort of implicit knowledge about the constraints and rules of the motor system; and they code information in a manner that considers future actions that will come from adopting a particular motor plan. In a similar way to P-intentions being directly related to D-intentions, M-intentions get their goals from relevant P-intentions. This overall process is therefore a kind of content-cascade, and we cannot fully understand any level's content without at least indirect reference to the content of

[2] Chiara Brozzo (2017) has argued that M-intentions fail to meet the core features of intentions, i.e. that they ought to least in principle be accessible to consciousness, have rationality constraints, and respect a consistency requirement. While these are serious concerns, which are at least partially addressed in Mylopoulos and Pacherie (2018), I will assume that M-intentions are a type of intention for my current analysis.

the other levels. Finally, M-intentions play a particular role in selecting relevant motor plans and then guiding the fine-grained motor responses required to complete the action at hand.

For an example of these three kinds of intentions operating for an individual jazz musician, consider someone deciding to perform with a jazz quartet on Saturday night. The moment that the decision is made, there has been at least one D-intention formed, namely, "I shall perform in this jazz quartet on Saturday night." This intention is reason-responsive - I can ask the musician questions about their intention to perform and call on them to provide further reasons for their actions. It also opens and forecloses decisions in light of satisfaction conditions, which will guide the musician's actions in clear ways, assuming that they are rational, and will further call on them to respond to criticism for failing to follow through on their plans or promises. For instance, in light of this particular D-intention, it would make sense to restring a guitar or practice a horn. It would not make sense to pawn off their instrument, get on a bus to leave town Saturday afternoon, and never be heard from again.

Skipping ahead to the day of the performance, the musician now meets up with the quartet and plays a gig. During the performance, P-intentions and M-intentions are developed and used. The content of the original D-intention is also made more specific; it is not just an intention to "play Saturday night" anymore, but an intention to play this song at this particular time. P-intentions concern certain decisions made during the implementation of the original D-intention in this case. It doesn't matter for the D-intention, especially at the time of original conception, whether or not the musician drives or walks to get to the gig. But they will eventually have to figure out some acceptable mode of transportation to get to the performance space. P-intentions may also include spontaneous actions that occur outside the purview of any D-intentions. The soloist who is responding to a new idea from the bassist, in contrast to the

soloist who is only anchoring their musical decisions in the harmonic structure of the lead chart plus previous schemas for how to play the solo, are examples of this difference.

M-intentions for the performing musician include elements such as fingering patterns, eye gaze, and low-level motor activities that allow for the successful completion of the performance. They generally occur beneath the level of explicit conscious awareness, and become more varied and fine-grained in the case of expert performers. In terms of how to relate P- and M-intentions, the musician may have a P-intention to start or to continue playing the song *Salt Peanuts*, but the M-intentions they use to execute that intention may be varied and sometimes occur without any explicit conscious control from the performer in question.

Although these sorts of intentions are relevant for describing how an individual goes about engaging with a jazz performance, they do not give a full story how such a performance is possible. To provide one, we must understand how a collective can come together to perform, in addition to how an individual may perform on their own or as part of a group. In order to further flesh out this account, we shall now turn to shared intentions.

Shared Intentions

We have thus far considered the nature of intentions for individuals. However, bebop performance is inherently a group activity, and it is not enough to explain the phenomenon by appealing to the intentions of musicians in isolation. Even if one wishes to reject a robust sense of shared intentions in this case, at least in the sense of those attributed to a group mind or group agent somehow metaphysically distinct from the individual performers, there is at least some kind of shared intentions at play. It may furthermore be the case that these shared intentions are not reducible to the structure of individual intentions discussed in the previous section.

For starters, if we look at the collective, the shared intention may have something like the

form “We intend to perform a song, X,” in contrast to pooling together individual intentions “I intend to perform a song, X”, just like the intentions of a dance group to perform in a park would also include some kind of group intention, or at least some sort of belief that identifies them as part of a group and acting as group members, and that separates them from the case of a mob running for shelter from the rain.

An additional way for motivating the difference between individual and shared intentions is the following thought experiment. Assume we are witnessing two different performances of bebop musicians playing *Blue Monk.* They agree in both cases that each soloist will take two choruses for their respective solos and maintain the same tempo throughout. In the first performance, the musicians jam the tune together in a rehearsal space. After calling out the song and counting off, they play from start to finish, including their agreed upon solo plans. In the second context, the same musicians are isolated within several soundproof, individual practice rooms.[3] An intercom is used to call the tune and stays on for the count off, but is shut off once the song starts so no auditory feedback is available. A metronome keeps the beat steady throughout the performance and the musicians stick to the agreed-upon solo plan, as before. We are then faced with the question are there any essential differences between these two cases?

I suggest that the result from these two performances would not only be vastly different in terms of quality and output, but also tells us something essential about the performance process itself. This difference becomes more acute if we consider the same idea applied to a Western classical string quartet. We can therefore extend the thought experiment a bit more,

[3] There are similar thought experiments vis-à-vis isolated practice spaces and jazz performance in independent unpublished work by Mark Risjord and Eric Lewis. Their use of the example, however, raises different points than the one on collective intentions I am focusing on here.

applying it both in the context of a string quartet playing a song together in one room and playing the same song after being isolated in individual practice rooms.

Still another step would be to record each of the four respective performances: one-room jazz rehearsal, one-room string rehearsal, jazz solo rooms, and string solo rooms. We then mix them together to form four respective finished recordings. The results from the one-room rehearsal rooms would be similar, at least insofar as they would both be coherent versions of the respective songs. The same cannot be said for the solo rooms, however. For instance, it will be easy to take the four individual pieces of the string quartet and fuse them together to form a finished song, albeit a bit less polished than usual. The recording of the jazz case, regardless of the skill level of the musicians, would be a disjointed mess. While there are important ways in which playing together has ramifications in both the jazz trio and the string quartet contexts (cf. Høffding, 2019), at the very least since some performance nuance would be lacking across solo room cases, the isolated performance of *Blue Monk* would fail to cohere in any meaningful fashion, instead of just lacking in performance nuance - it would devolve instead into three people playing on their own accord instead of performing as a disjointed trio.

One thing that is missing between the rehearsal rooms and the same room contexts would be the presence of shared intentions-in-action. What makes intentions-in-action, or any other kind of intention, shared in the first place? One possible way for becoming shared is that we simply pool together the individual intentions of the musicians that originally had the form "I intend to perform a song, X." The shared intention would then be nothing over and above the concatenation of individual intentions of that form. Such an approach would be akin to what is known as the summative view (Quinton, 1975).

Applied to a case of beliefs, for instance, a summativist would say that the group belief

that A can be fully cashed out as all (or most) of the (relevant) group members having the individual belief that A. There are several well-documented issues with a summative view (Gilbert 1987; 1994), especially in regards to claims about group beliefs or desires. At the same time, there is at least one additional position that is similar to summativism that I will grapple with below. Both of these accounts are linked in rejecting the existence of shared intentions as relevant for underlying ensemble performance, although they do so for different reasons.

The other position is methodological atomism (Epstein, 2018), which I shall interchangeably refer to as atomism for short. Defenders of atomism suggest that we should understand any purportedly group phenomenon, including the nature of improvisation during an ensemble performance, as not in fact a properly group activity, but rather an individual activity that takes place in a group context. It also takes people to be fundamentally individuals, rather than inherently social creatures. For example, consider the case of a staff at a professional kitchen making a meal. While proponents on all sides of the debate grant that the final meal results from a group effort, according to atomists, we shouldn't claim that there is some inherently group-like kind of activity going on, beyond perhaps the fact that several people were involved with making different parts of the dish.

Combining the main assumptions from summativism and atomism leads us to a particular version of the social contract view regarding human behavior and our explanations of it (see Hagberg, 2016). This view entails additional metaphysical assumptions beyond the nature of individual vs. collective mental states. A proponent of the social contract view holds that our mental states are unobservable, internal phenomena, for example. They likewise hold that only individual organisms can have a self or be agents. Although one could accept either

summativism or atomism without agreeing to all of these additional assumptions, the social contract view would fall if one of them were to fail.

In terms of benefits from the social contract view, following Hagberg (2016) and with contentious ontological assumptions notwithstanding, I grant that it may do justice to certain musical phenomena. It will do especially well for genres that have a heavy amount of preparation before the time of performance, such as in the case of a symphony. Whether or not such a view is ultimately satisfactory in the case of bebop jazz, however, is not as clear. In order to consider the jazz case in more detail, I will shift slightly from the social contract view to focus on methodological atomism. In turn, since atomism is an essential feature of the social contract view, its problems will similarly be problems for the latter position. As a result, a need to reject methodological atomism can be taken as a need to reject the social contract view.

Methodological Solipsism

To begin our considerations on methodological solipsism, we can consider an argument in favor of an atomist view from Stefan Carris Love (2016). While bebop is a style of performance where all performers are simultaneously improvising, as noted in the introduction, it is not the case that every performer must be responding in an egalitarian manner to all possible options presented by their co-performers. As such, while there are important group aspects to the negotiation and development of a bebop performance, there are *prima facie* reasons for holding that the individual musicians are able to (individually) make all the relevant decisions within a given performance.

Love explicitly considers three reasons why improvising jazz ensembles are ultimately explainable as individual acts, even if they take place during a group context (2016, p. 71-2). First, the soloist takes the position as leader of the band during their solo, and thus becomes the

center of attention. Second, all of the constraints provided by musicians comping during the solo will, at best, act as local constraints. The soloist is not required to act on these constraints if she or he does not wish. While some musicians draw inspiration from what the other ensemble members played, others demand that the comping musicians stay out of their way. Third, Love suggests that solo jazz improvisation is in principle the same as soloing in a group context, especially if we consider the case of the jazz drummer who often solos without any other musical accompaniment even during ensemble performances.

Love further suggests that these three points can be clearly contrasted with the case of improv-theatre. While there may be times when one person is clearly the focal point of attention in theatre, it is often the case that everyone is equally important to watch. Each actor is also heavily constrained by the choices made be their fellow actors. If you start the scene by greeting me as John the librarian, I must either uptake that role or at least engage with it directly before moving past it. Likewise, to harken back to the prat fall example, if one of the actors falls during the scene, everyone *must* respond. Finally, while it is an open question whether or not the cognitive and performance processes in individual cases of improv-theatre are identical to the group cases, it remains certain that there are options either available or foreclosed in the latter case that are not present in the former, as seen with the first two points here.

While I grant that these considerations are important distinctions among theater and jazz, I suggest that Love is mistaken on his position that jazz solos in ensemble performance are not inherently group-based and shared phenomenon. First, while it is true in many cases that the soloist is the leader, at least for a time, it is never the case that they can act independently from the rest of the band. This point is most salient in regard to the connection between the soloist and the rhythm section. In addition, just as the soloist can ignore the prompts and prodding from the

band, other members can ignore the will of the soloist. Choosing to do so may have negative professional ramifications in either case, yet there are also times when this sort of process pushes creativity in interesting and important ways.

Finally, there are additional grounds to reject the third point that soloing as an individual is identical to soloing as part of a group during jazz performance. I grant that there are certain similarities when soloing by oneself and when soloing as part of an ensemble, and we may find that such similarities entail similar psychological and neurological processes. Yet these similarities are not such that both contexts are *essentially* the same. fMRI studies furthermore put pressure on a strong similarity claim; for instance, there have been different neurological patterns found in individual vs. collective contexts, seen most clearly in the DLPFC having decreased activation in solo improvisation and increased activation in group improvisation (see Beaty, 2015, p. 113). Being in a band provides stimuli and constraints that are simply not there in the case of solo performance, in short.

One may here ask would it be possible to just replace the band with a recording? While one could receive unique stimuli with a prerecorded track, there is something different between that case and the case of playing with other musicians. This point becomes more acute when we consider the case of the audience perspective. Whether a soloist ultimately decides to ignore, include, or play alongside the material provided by the rest of the band, the audience is always hearing those choices within a particular musical context that may not have been performed in that particular way, at least during an ensemble performance. The song as performed in the group is therefore thicker than just the individual solos and a backdrop of music. I further suggest that we can hold this claim even in the case of drummers performing in a group, and even though they often do solo unaccompanied in ensemble performances. This point is especially acute in

bebop drumming.

There are two main ways that a drummer plays a solo in bebop. The first way involves "trading fours," where the drummer and another instrumentalist take a series of quick, alternating solos lasting four measures each. Even though the drummer is playing unaccompanied during their respective part, they are tied directly into the choices of the other soloist in an exchange and development of ideas. With the fast speeds and short lengths of these solos, most are constructed as a sort of conversation with each musician drawing on the ideas raised previously and developing them accordingly.

The second way involves taking a solo chorus of the same length as the others, usually as the last solo before the outro. Even in cases where the drummer may be performing a full 16 or 32 measure solo by themselves, however, they are doing it in the context of the song that they are playing. While being an unpitched instrument means that their musical sounds will operate differently than the notes played by pitched instruments, they are making their choices, again, in reference and relation to the song currently being performed by the group, and even if some of these performances are more linked into the group performance than others. Based on these considerations, I suggest that Love's atomist account will ultimately fail to capture the essential structure of intentions at play during jazz ensemble performance.

An additional issue here is that atomism fails to offer an account for the ways in which the intentions between different musicians are *connected* in the moment of performance. Such connected responsiveness requires that individual musicians constantly monitor the developing soundscape during performance in light of their own interest to play a song. Nevertheless, it also requires them to respond to what is going on in the music outside of their control and in relation to their co-performer's respective goals, interests, intentions, and desires in a myriad of ways.

Even the soloist who prefers sparse accompaniment is still having their intentions constantly negotiated in relation to their co-performers.

At this point, however, one may claim that the situation I've presented so far, especially in regards to the thought experiment, is either too abstract or that I have only shown something about the performance context and joint attention, rather than something about the nature of shared intentions. Both atomists and summativists would attribute nothing else over and above the concatenation of individual intentions in the form of "I intend to play *Blue Monk*," perhaps also allowing that band members do so with the knowledge that others in the group have similar individual intentions. They would also hold that any beliefs about what ought to be played can be exhaustively realized by the combination of individual beliefs on what to play next. In response, I suggest that if I were only constrained by this kind of intention, there would be little to explain the sorts of negotiations that are ultimately present during the unfolding of a musical performance. I will return again to the question of negotiation in the section on Bratman's theory applied to jazz improvisation below.

Another concern with my replies to Love can be extended the practice room thought experiment as a ground for rejecting atomism. Everyone from the staunch individualist to the group mind proponent would grant that there is something distinct in the case of performing a song together and performing one's part of the song on their own (*pace* Love's third point about jazz improvisation considered above). Thus, this concern would be that my thought experiment at best begs the question against the atomist, if not completely missing the mark of their worry. More evidence may therefore be needed to be establish why these accounts don't work in the case of ensemble bebop improvisation.

Turning from thought experiments to actual experiments, a further point in favor of the

summativist and atomist positions comes from recent work by François Pachet, Pierre Roy, and Raphaël Foulon (2017). In this particular study, the authors track what they refer to as the "score effect" in collective jazz improvisation. This effect concerns whether or, perhaps more accurately, how much an improvisation can be attributed to the link between musicians and the score/lead sheet, instead of the engagement between musicians during performance. The latter engagements, which the authors refer to as "content-based interaction," covers more robust exchanges among musicians that are often thought to be an important part of jazz performance. Content-based interactions likewise move beyond mere synchronization to include high levels of performance, such as sharing melodic, harmonic, and rhythmic components. The score effect covers any interaction between musicians that is mediated through the lead sheet, which in bebop constrains the harmonic choices of both soloists and accompaniment.

Starting from these two kinds of interaction, Pachet et al. attempt to establish the extent to which choices made during a performance are due to the score effect. They addressed this question by comparing the relation between two instrumentalists performing together, on one hand, and when they played at different moments in time but with the same score, on the other. The data for their analysis was drawn from a series of performances by the Mark d'Inverno Quintet, a bebop band, recorded both live and in the studio. The authors specifically focused on the relation between the sax player and the drummer and bassist.

From within a given song and cluster of musical features, the authors computed both the total interaction and the amount of score effect via a set of interdependency indicators. Their results indicated the majority of content-based interaction could be explained away by statistical noise instead of being an important part of bebop performance, and they conclude "either

content-based interaction in jazz is a myth or that (more probably) interactions do take place but at yet unknown musical dimensions" (2017, p. 174).

Taken at face value, these conclusions push against my position. However, I believe that we can respond to them in a rather straightforward manner, especially if we consider more about the "unknown musical dimension." One clear candidate is to focus on higher levels of musical organization instead of the features isolated for analysis. Considering these levels would likely push us away from atomism towards at least some sort of minimal or maximal account of shared intentions. Another option would be to dig deeper into the specifics of the embodied interaction that take place during performance. Especially considering my immediate focus on embodied music cognition, I suggest that this move may provide evidence of more robust content-based interaction among musicians. Finally, the authors note that their current study did not gather data to consider either of these options (2017, p. 174). As such, while these open possibilities are not yet part of a knockdown counterargument, the possibility of finding more interaction in higher-level musical features or in embodied interactions should give pause to the atomist who wishes to claim this study as strong support for their position.

In addition, the authors presuppose that content-interaction and the score effect should be viewed as mutually exclusive options. I suggest, *contra* this assumption, that the creation of the music itself remains a robust interactive phenomenon that happens to involve both the score and fellow musicians. Sometimes individual performers may decide to interact more or less directly with each other. In either case, the origins of the performance simply cannot be explained solely by appeal to what is included in the score. The full causal story of a given performance likewise cannot be reduced to the interactions between individual performers and the score at hand. The

performance is, in other words, nonreducible to the heads and actions of individuals and returns us to consider the kinds of shared intentions that may be relevant for bebop improvisation.

Shared Intentionality in Bebop

To understand shared intentionality in bebop, I first return to expand on Hagberg's point about an irreducible group component to jazz improvisation. To develop a positive account, Hagberg (2017) follows a "subtractive method," citing Ludwig Wittgenstein and Arthur Danto as others who have utilized this general approach. The subtractive method starts with a jazz performance, strips away individual intentions, finds what's left, then turns to explain the remainder. In the case of ensemble bebop improvisation, an important part of the remainder comes from the fact that no individual musician can control the collective output of a song. To address the remainder, Hagberg (2016; 2017) has offered several analyses that offer an illuminating account and draw on the work of Bratman, Searle, and Gilbert. I will return to various elements from these analyses below.

An additional question remains of how best to conceptualize the results of this subtractive method. A strong reading is that we should understand an improvising group as an entity standing above the individual musicians making it up. If we take this entity to have some independent ontological standing, we may postulate particular cognitive and mental states to it as well. An alternative, weaker reading is not to posit some sort of group agent but, instead, note that the group improvisation provides important constraints and opportunities for the individual performers. These constraints, however, would be more robust than mere suggestions for options to follow. We can see something like the breadth of this constraint-based analysis in Adam Linson and Eric Clarke (2017), where, in response to hearing a series of high notes being simultaneously performed, they ask:

> As a participant, what is significant to you about this state of affairs, and which particular aspects are you attending to? Simply that there are many high notes? The specific collection of pitches, or the rhythm? That there is an audible resemblance to animal sounds…[o]r that it makes reference to a musical style and its historical period or social milieu?...Meanwhile, the attention and interpretation of every other improvisor on the ensemble guides and is guided by each player's opportunities for action in the situation, shaped by the affordances of the instrument, the player's liveliness of fatigue, his or her relative social confidence or familiarity with the group…and so on. (p. 64).

All of these proposals can be further situated in light of Alan Carter's (1990) distinction between individualism, interrelationism, and collectivism. The summativist and atomist views, for reasons explored above, are examples of individualism. They are likewise not fit to understand shared intentions in jazz improvisation. Supporting the collectivist view would entail accepting the robust reading in the previous paragraph and positing that the ensemble output can be attributed to some kind of group mind or group agent. While I believe that this route has promise in the case of jazz improvisation, and we shall consider one kind of collectivist option in more detail below, interrelationism remains the main live option that will likewise be explored more below, especially in the case of Bratman and the extended DPM model.

In addition, while Hagberg offers an important series of considerations in defense of shared intentions for jazz performance, one potential problem with drawing on a constellation of Searle, Gilbert, and Bratman is that doing so glosses over fundamental conceptual differences between their approaches. On one hand, I grant that these different accounts may serve useful for explaining different parts of jazz improvisation, as Hagberg shows and as we shall soon explore

in additional detail. On the other hand, further analysis is required to see just how well these respective theories work together. Such additional analysis is likewise important to find the fundamental cognitive machinery that is needed to understand jazz improvisation. Plus, considering an important role for parsimony, I suggest that our final account should be as minimal as possible to explain the phenomenon of bebop improvisation at hand.

One of the questions flagged at the outset of this chapter concerned whether maximalist or minimalist accounts of joint action are better suited to capture bebop improvisation. In what follows, I focus on three different possibilities for capturing the joint action underlying bebop performance. The first is Gilbert's (1989; 1990; 2007; 2014) joint commitment account, which entails the idea that a joint action occurs when a group jointly commits to act as a body. The second is the planning theory of shared agency developed by Bratman (1987; 1993; 1999; 2014), which entails that a joint action requires a particular structure of interdependent individual intentions that mesh properly and are taken up under common knowledge. The third account is the extended DPM account (Tollefsen, 2014; Pacherie, 2013) which takes the basics of the DPM model introduced in the individual intentions section and extends it to joint action.

All three accounts will ultimately have something important to say about the phenomenon of jazz improvisation. Among them, however, I suggest that only the extended DPM account provides the grounds for explaining the entire spectrum of bebop improvisation, from high school jazz bands to the best jazz bands in the world. By "entirety", I mean that the extended DPM model presents a common structure of shared intentions as they unfold during all kinds of bebop performance from novices to experts. I will also consider whether or not the extended DPM model could be made even more minimal by comparing it to an enactivist alternative. Although I conclude this chapter by holding that we don't need to go more

minimalist than the DPM model, the next chapter will consider another way of approaching bebop jazz that is closer in spirit to enactivism and shows another alternative kind of account to understanding jazz improvisation.

Gilbert's Joint Commitment: A Maximal Maximalist Account

For the past several decades, Margaret Gilbert has developed an account of shared agency around the idea of joint commitment. An essential feature of joint commitments is that they are not reducible to individual commitments. Instead, they are produced when individuals come together in cases of joint action and shared agency. In earlier iterations of her account, joint commitments were attributed to a plural subject. While Gilbert was ontologically permissive as to the metaphysics of plural subjects,[4] the core of the concept stands conceptually separate from a mere pooling of individual intentions. Indeed, this feature is one of the major differences with the other two options we will consider below.

Following a commonly used example, consider a joint commitment between Emilia and Mike to take a walk *together*. While perhaps being explicitly stated - assuming Mike asks if Emilia wants to go for a walk and she responds "Of course!" - it is possible for a joint commitment to appear without being so, such as if the pair happened to meet at an intersection and continues to walk home together. In a similar manner, it is possible for the joint commitment to be made without anyone being explicitly aware about the nature of joint commitments as a conceptual apparatus. Joint commitments can thus be made even if neither Mike nor Emilia has studied philosophy nor come across the literature on joint action and joint commitment. They therefore hold for basic human social actions, not only for the social actions of those steeped in the philosophy of action.

[4] While I will discuss plural subjects more below, Gilbert has moved away from the term in more recent work.

Joint commitments furthermore entail a "package of rights" that opens up a series of expectations within a group. Emilia would expect Mike to meet her at the agreed time and place for their walk and Mike would expect Emilia to walk at a reasonable pace. These rights correspond to the possibility for giving a rebuke if expectations are violated without the joint commitment first being dissolved or resolved. If Mike suddenly starts to speed walk, Emilia would have the right to stop him or at least call on him to explain why he is running ahead.

This notion of rights only speaks to one important part of this concept, however. Another important part is a plural subject, which can itself be the bearer of mental states over and above the individuals who make it up. One may here ask why posit mental states to a plural subject in the first place? For one, some shared mental states may only be realized in light of institutional and material conditions. Moreover, there may be cases where the same individuals make up the full membership of different groups, but the groups *qua* groups would have vastly different beliefs about the world. For instance, we would attribute beliefs about the best new books the library ought to buy to the library book committee but not the department hiring committee, even if the committee members are identical in both cases (cf. Gilbert, 1987). Other times, mental state attributions may hold for group cases from functional similarity, such as in transactive memory systems (Sutton, Harris, Keil, & Barnier, 2010), where memory is dynamically distributed across the collective in ways that cannot be limited to any individual member and often requires active engagement as part of creation, encoding, and retrieval of memories.

This ability for a plural subject to be the bearer of distinct mental states offers an explanatory tool to address the situation as described by Linson and Clarke above concerning how an individual performer will engage with aspects of performance. There are expectations and possibilities left open during a performance. The specific developments of a given ensemble,

through processes of agreement and disagreement, supporting and blocking, pushing and pulling, will develop without any individual member controlling the entirety of the process. How audience members choose to interpret the meaning of what is played will also be related to the plural subject of the group, and certain interpretations will lend themselves more immediately for one group in contrast to another. As a result, some of the responses to the environmental situation will be driven by the individual, while others will be driven by the ensemble itself.

From this point in particular, we can move our analysis from walking together, deliberating together, and remembering together to the case of performing together. Hagberg's (2017) application of Gilbert's account to jazz performance takes the insight of the plural subject as a starting point. He then notes that an appeal to it is useful to account for how jazz improvisers talk about the music taking over and being in situations where the group is going where it was best to go, especially for those that have developed a strong ensemble identity over time.

To elaborate on these points, Hagberg considers a 1987 performance of "On Green Dolphin Street" by Stan Getz's quartet. Moving through the song in detail, he notes an important series of transitions that separate the preconditions of a great performance – for instance, at the 32 second mark, "we have the highly advanced musical version of Jack and Sue walking together, indeed, walking together perfectly—*and no more*" (2017, p. 307 emphasis added) – to the realization of a great performance later in the song, especially at the start of the first improvised sax solo. From the speed and fluidity of the changes between the instruments, Hagberg suggests that "such a musical event *could* not be thought or enacted by one person" (2017, p. 309, emphasis in original). Gilbert's plural subject is thus realized in action for the Getz quartet.

One may still wonder whether or not it is conceivable that a performance such as this one

is enacted by a group of individuals rather than a plural subject. To that point, it will help to quote Hagberg in full (2017):

> In this sense, one argues back from the result: if ensembles achieve the results they do, and that we have empirically before us as listeners and phenomenologically in the experience of players, the irreducible 'remainder' that makes the difference between a performance at the height of this art form and one that, shall we say, aspires to such heights, then the concept of the plural subject underwrites the very possibility of this music. (p. 305)

For Hagberg, the high moments of jazz performance either entail or result in the creation of a plural subject, and the burden of proof is on the individualists to show that this outcome is not the case. If Hagberg is correct, the nature of joint commitments made by highly trained musicians such as we see in this situation points to how a plural subject can be constituted in expert performance and is a necessary feature of describing jazz improvisation.

Since we now have a case of a plural subject realized in a successful expert performance, the question becomes whether we need this robust level of joint commitment to account for all successful instances of jazz performance. I believe that we do not for several reasons. First, Hagberg clearly notes that parts of his analysis draw from the feelings and thoughts of expert musicians. We also see in the quote above that his argument rests on the fact that we make a distinction between a performance that reaches the level of a specific Getz quartet performance from cases where that goal is unreached. This unreached part requires considering if a plural subject is a necessary feature to explain how musicians play together during, e.g., an open jam.

On one hand, there are certain requirements that any musician who takes part in an open jam must follow. These may be understood as a type of joint commitment among performers,

similar to what occurred in the case of Mike and Emilia taking a walk together. For instance, anyone wishing to take part in a jam will adhere to certain stylistic and genre norms to the best of their ability. This requirement is particularly acute in cases with a demanding bandleader running the show or if there is a presupposed minimal level of musicianship required to participate. A clear example of a demanding leader on the bandstand was tenor saxophone player Eddie "Lockjaw" Davis, who noted of his experience as a bandleader that (Taylor, 1993):

> You're trying to play, but you're also trying to police the bandstand so that nobody is in the way. There were times when it almost became violent. One or two individuals felt I had no right to make the decision…In some instances I was labeled a tyrant, but on the whole the guys appreciated it. (p. 44).

Even in cases where the jam does not include a musician of Lockjaw's status, however, there are certain joint commitments at play. The need for these commitments may be one important reason behind why Hagberg claims a plural subject underlies "the very possibility of this type of music," instead of merely holding that the plural subject underlies the pinnacle of performances.

Nevertheless, while this sort of response may establish a role for some kind of joint commitment in all cases of jazz performance, it is not clear exactly what the content of these joint commitments are in the case of jams. So far, I have made explicit reference to some variation on "we shall play this song together" as the content of shared intentions and, by extension, joint commitments. While such a case may be common for bands who have taken time to rehearse beforehand or groups with members who are knowledgeable in the history of jazz, it is an equally common experience in open jams that some, and perhaps even most, of the performers will have no pre-existent knowledge of a song that is played. In these cases, while

there may be a collective agreement to follow certain norms and conventions, it's not clear that the performers are tacitly or explicitly agreeing to play the song as a body, unit, or group.

Gilbert may here remind us that one does not need to understand the notion of a joint commitment to operate under one. However, the problem we are currently facing runs deeper than the fact that the performers at a jam don't think of themselves as performing as a group. The issue is rather that the successful performance need not include an irreducible plural subject in order to occur. It may merely require a concatenation of individual intentions with the proper sort of attention and attunement to others and the music in order to navigate the performance. This result puts pressure on the need to have a plural subject as a necessary conceptual apparatus for our explanation, yet it does so without turning us back into the fold of a traditional, anti-group summativist account. It also sets the stage well for turning to the planning theory of agency, which takes intentions as closer to the concatenation of intentions with the proper sorts of common knowledge and meshing.

Before switching to Bratman's account however, and while the above considerations may put pressure on the need for a plural subject as a necessary feature of bebop jazz improvisation, I wish to stress that there still is an important role to play for joint commitment. There are two additional options to establish this point. First, following John Michael, Natalie Sebanz, and Günther Knoblich (2016), we could leverage considerations about a weaker sense of joint commitment that supports the development of the full-blown kinds in Gilbert's account. This minimal notion of commitment, taken by Michael et al. in a psychological sense rather than the full-blown normative sense, consists of two features (2016, p. 4-5): (1) There is an outcome that an individual agent either desires or has set as the goal of their current action and (2) There is a contribution from a second agent that is crucial to bringing about that desire and/or goal. The

relation between (1) and (2) can be supported by a combination of emotions and sentiments, as well as goal slippage and expectation fulfillment.

Second, we could focus on the normative commitments that underlie jazz performance, regardless of whether or not the success conditions of an excellent performance are met. If we follow this approach, we would say that there is a joint commitment to certain performance practices and norms, regardless if one is in an open jam or one is expert among experts, and they are jointly established for a successful performance. A particularly important notion in this regard comes from the role of trust (Wilson & MacDonald, 2017; MacDonald & Wilson, 2006). Trusting one's co-performers allows musicians to take risks they otherwise would not, which heightens the possibility of a better performance among the group. It also means that individuals will be able to respond in ways more characteristic of a plural subject, since the sorts of responses are geared towards the collective.

Yet, as we saw previously, the grounds of these two senses of joint commitment need not be some kind of irreducible plural subject, and it remains the case that these performance norms may be cashed out as the concatenation of individual norms, instead of being attributed to a group acting together as a body. With that in mind, we now turn to Bratman's account to see if his planning theory can explain other aspects of shared intentions in jazz improvisation.

Bratman's Planning Theory: A More Minimal Maximalist Approach

According to Michael Bratman (2014), something like bebop is the kind of phenomenon meant to be captured by his account. In his words:

> [M]y focus will be primarily on shared intentional activities of small, adult groups in the absence of asymmetric authority relations within those groups, and in which the individuals who are participants remain constant over time…*My*

> *interest will be primarily with duets and quartets rather than symphony orchestras with conductors*… (p. 7, emphasis added).

As a result, the extended use of bebop jazz as a case study is essential to evaluating Bratman's theory. In addition, Bratman's account, while not in the minimalist camp, is one of the more minimalist alternatives within the maximalist approaches. For instance, in contrast to Gilbert's account and the irreducible concept of a plural subject, Bratman rejects the idea that we need additional tools to capture the phenomenon of shared agency above and beyond individual intentions taken in distinct ways.

What exactly are these distinct ways of taking individual and shared intentions? According to Bratman, intentions are fundamentally *planning* states. The structure of planning, in addition, explains the difference between belief-desire constellations and intentions flagged at the start of this chapter. It also explains how intentions are integrally related to motivating and guiding actions. The idea that intentions are plan states has also received a general level of acceptance within the philosophy of action literature. With that being said, the senses of planning invoked by these theories are different from the prototypical or folk notion of plans that are more familiar to most people.

One example of the folk notion of plan would be a shopping list. For starters, a shopping list includes a particular set of considerations about what I plan to buy at the store. Buying things at the store not on the list would be unplanned actions, in contrast. Beth Preston further explicates the prototypical notion of plans through the inclusion of four criteria: they have a specified end; there are specific steps for achieving that end; there is a formulation of both steps and ends ahead of the action; and the making of this plan was a deliberate, conscious, explicit process. (2012, p. 45). While there is leeway regarding whether or not all of these criteria are

necessary or sufficient for plans, they offer a good starting place to consider what counts as a prototypical, folk plan.

While folk plans can exist without preconfiguring everything one must do to complete them in advance – the shopping list remains a plan even if it doesn't include a map for walking around the store, after all – they are more substantive than the sorts of plans that Bratman has in mind for his planning theory of intentions. Intentions are more like partial plans, especially if we compare them to the full folk sense of plan. Yet they also maintain other features that are similar to folk plans, especially when it comes to guiding actions.

As planning states, intentions have normative aspects, such as means-end coherence and providing motivations for action. We can recall here the individual intention to get a beer, for example. Having this intention, whether or not it is stated aloud, results in me taking certain actions in order to reach my goal of getting a beer. While it may be possible to have an intention without acting on it for an extended period of time, there will eventually come a time where I either act on that intention or, if not, I open myself up to questions why I failed to act. We can see a similar basis of these elements in the shopping list. If my Mom sends me to the store with a list and I return with nothing from it, I will have failed the normative commitments inherent with the plans embedded in the list and my use of it while shopping. These commitments are also themselves part of the motivation for action, since I decide to buy bread, milk, and cheese because they are on the shopping list.

Both shared and individual intentions are of the same metaphysical kind for Bratman (2014). Shared intentions are planning states, as a result. The difference between individual and shared intentions is instead found in two other areas. First, for their content, since the aim of the shared intention is something that *we* intend, it tracks a different level than the individual sets of

actions I utilize to reach that goal. Second, there is particular kind of interrelation among individual intentions to form shared intentions. The collective case further contains five distinct features: intention conditions; belief conditions; an interdependence condition; a common knowledge condition; and a mutual responsiveness condition.

To flesh out this structure in some detail, let's consider the Miles Davis quintet developing the shared intention to perform *Freddie Freeloader*. For the (1) intention condition, each member must have the intention of the form "I intend that we J [i.e. perform *Freddie Freeloader*]", and each of these intentions must be supported by the fact that others have that same intention. For the (2) belief condition, each member must actually believe that the other member holds the same intention. Perhaps the band could successfully perform *Freddie Freeloader* without these sorts of beliefs, but it would be a case of mistaken luck rather than a case of a shared intention to perform the song. For the (3) interdependence condition, the continuation of each band members' intentions rests in part on both their beliefs about the other members' intentions as well as the fact of the matter that those intentions are present. For the (4) common knowledge condition, all of the above must be common knowledge among the band. Finally, for the (5) mutual responsiveness condition, the individual players must show their willingness to take part and support the shared intention to perform as the performance unfolds.

There are several criticisms that have been leveraged at Bratman's account of shared agency in general. One is the fact that it seems unnecessarily cognitively complex, especially in relation to the successful joint actions performed by children (Tollefsen, 2005) and non-human animals. Another is that it doesn't do well in accounting for present intentions. These are major reasons why Bratman's account is not well suited to provide the entire story of shared intentions for jazz improvisation and I will return to the second worry below.

There is an even more pressing point of concern for using Bratman's account to explain the shared agency behind bebop performance, however, regarding the definition of improvisation as an unplanned activity. If this concern holds, there will be grounds for rejecting the possibility of using any of Bratman's theory to understand jazz improvisation. Specifically, the worry is that the focus on a planning theory is at odds with improvisation by definition, especially if planning is tied into making some kind of preset plan.

If we hold an inherent tension between plans and improvisation, there is an immediate problem with Bratman's account of shared intentions being applied to the case of improvisation. Preston has a detailed version of this worry in her book, and I will turn to her arguments in a moment (Preston 2012, chap 3). Before doing so, there is one possible quick reply to this tension that needs to be considered and rejected. Doing so further helps establish the exact stakes of Preston's more detailed concerns with a planning theory of intentions in improvisation.

The quick reply is that intentions are best taken as *partial* plans, and partial plans would fit well into the spectrum of improvisation I considered in chapter 1, In other words, this reply posits that plans by themselves are not always antithetical to improvisation. It is only the idea that all intentions are preset plans that would pose a problem, and insofar as Bratman accepts both future and present intentions, he clearly leaves open a sense of intentions that does not collapse into preset planning. At the same time, although Bratman's notion of plan is distinct from more robust notions of planning, there are remaining ways in which even his more minimalist notion of plan may be problematic in ensemble jazz performance.

Consider talking to a group of jazz musicians after witnessing a particularly powerful performance. During this conversation, the guitarist tells you that they had performed their solo exactly like they planned or intended before the show (i.e. in Bratmanian terms, once each

individual musician formed the intention "I intend that we J"). One would likely be taken aback by such a comment. After all, if the musician had planned everything beforehand, they wouldn't be taking part in an improvised activity in the first place. While we would likewise be taken aback if the musician said the entire performance was done without any guidance or knowledge of what was being performed at the time,[5] there is something distinct about the problem of claiming that an improvised performance could be planned ahead of time.

Bratman may respond that this worry still rests on a confusion between the notion of plans in the planning theory and full-blown folk plans. In addition, shared intentions are meant to be the grounds of bargaining for the overall activity. Thus, the previous example of the guitarist saying that the improvisation went according to plan, on Bratman's terms, may be better captured by rewording the claim into the form: "that solo respected the basic, tacit agreements we had about what is acceptable or not in performing the song". And, stated in this reformulated way, we again find that intentions as plans may not a problem for a Bratmanian account of improvisation.

However, the worries about plans have not been fully cleared yet. As Preston (2012) has noted, this reply fails to capture another variation on the worry that a planning theory of intentions is a problem. The important issue, here, is that intentions are *by definition* plan states on Bratman's account, while improvisation is by definition some kind of *un*planned activity. Insofar as improvisation is unplanned, on this definition of intentions, it will also be by definition unintentional. After all, if intentions are defined as plans, then improvisation is by definition a process that lacks intentions. Yet improvisation is clearly something that involves some kinds of intentions, which pressures the idea that intentions should be by definition plan states.

[5] This claim should be kept distinct from instances of deep absorption for the sake of the current analogy.

After leveraging the above critique, Preston explores one possible reconciliation of understanding improvisation along the lines of what Bratman calls "purposive cognition" instead of "planning cognition." Purposive cognition is a widespread phenomenon for non-human animals as well as infants and young children. It is also sometimes the mode in which otherwise full-blown planning agents may act. In contrast to planning cognition, purposive cognition is driven by a concatenation of beliefs and relevant pro-attitudes.[6] It is focused on goal-directed and consciously guided activity lacking the presence of full-fledged intentions. As a result, insofar as it lacks intentions, there is not the same sort of norm-based, reason responsive behavior at hand.

While going this route is one plausible alternative, I suggest that there are at least two additional replies that could be given while dissolving the tension and maintaining a planning theory. The first is to again reiterate that, on Bratman's account, intentions as planning states are best understood as *partial* plans. While Preston is correct that tethering intentions to plans will fail to capture important differences between planning and improvising, the vast majority of improvisation occurs with reasonable sets of preconditions and preexistent intentions. This point is a main reason why my general definition of improvisation separates strict preplanning and improvisation, but it does not say that improvisation is necessarily an unplanned activity.

The second reply, which is an extension of the first mixed with a reply to an earlier version of the current worry, highlights that the plans driving action provide a normative structure that helps guide bargaining what to play during the song. Some of this bargaining may be offloaded onto the song chart and the attendant structure of the song's harmonic features. However, the underdetermination of the score or lead sheet in jazz improvisation means that parts of the performance will only be fixed during the performance itself. The concatenation of

[6] These pairings may be taken in light of prior intentions. If that relation holds, then there is an overlap between planning and purposive cognition.

individual intentions along the five features above would ground the navigation of the performance as it unfolds over time, even if intentions are plan states.

As a result, we find that Bratman's planning account is not intrinsically opposed to an explanation of improvisation as such. Nevertheless, while these points show that we can have a planning theory of improvisation, they do not address the remaining question of what present intentions are in the case of improvisation. In particular, as Preston (2014) and Tollefsen (2014) both note in distinct contexts, Bratman's account has little to say about how shared agency and collective action works in the moment.

While bargaining structures and partial plans may get us deeper within the context of improvisation, these elements alone fail to explain *how* improvisation works in real time. In other words, while Bratman's account does a good job of explaining shared D-intentions, it does not offer us much in the way of understanding shared P-intentions (or M-intentions). Or, returning to the definitions of improvisation from the introduction, Bratman's account primarily explains important parts of what happens with improvisation in general, yet it doesn't speak to the more specific nuances of jazz improvisation as such. Realizing this sort of planning structure can therefore provide sufficient grounds for understanding how a group navigates important features of a general performance. It unfortunately has little to say about the ins and outs of jazz improvisation in particular.

A Minimalist Account: Shared Intentions on the DPM Model

As noted by John Michael (2017), the turn towards minimalist accounts offers several benefits for an account of joint action in music, most notably in regards to ecological validity. He furthermore flags three components that are relevant for minimalist accounts: representations, processes, and coordination smoothers (2017, p. 161). While we considered the first two areas

for the DPM model in individual action, we now turn to exploring all three in the context of joint action.

Most of the representational format in the DPM model will be the same across both individual and shared intentions, namely the conceptual, abstract format of D-intentions, the indexical, ostensive format of P-intentions, and the motoric, action-centered format of M-intentions. However, other representational elements will be different, especially in the case of P- and M-intentions. Deborah Tollefsen (2014), in particular, has developed an extended account of the DPM model into a dynamic theory of shared intentions, which I will turn to introduce in the following paragraph. In addition, since I accept that Bratman's account is satisfactory for considering most of D-intentions in jazz improvisation, I will take the primary focus to be on how the extended DPM captures shared P- and M-intentions for the remainder of this section.

After granting a broadly Bratmanian picture at the level of shared P-intentions – i.e. shared P-intentions will be combinations of interconnected, meshing intentions with common content – Tollefsen notes that they nevertheless receive some important modifications in the shared DPM model. First, the content of shared P-intention needs to track the action as a whole, in addition to one's individual contributions to the joint action. Second, the common knowledge requirement should be replaced (or at least augmented) by joint attention that allows the individuals within the joint action to be mutually aware of the intentions-in-action of the others.

An important difference between joint attention and common knowledge is that the latter requires more detailed, nested knowledge and complex belief attribution. To have common knowledge, I have to know that you know that I know something about X, and *vice versa*. In contrast, we can establish joint attention without these layers of knowledge and meta-knowledge.

Musicians playing together can jointly attend to the song that is being performed without A knowing that B knows that A knows that the song is being played. They simply play *together*.

An important source of work on minimal joint action comes from the research of Natalie Sebanz, Günther Knoblich, and their colleagues (Sebanz, Knoblich, & Prinz, 2003; Sebanz & Knoblich, 2009; Knoblich, Butterfill, & Sebanz, 2011). In a classic review paper, Natalie Sebanz, Harold Bekkering and Günther Knoblich define joint action as "any form of social interactions whereby two of more individuals coordinate their actions in space and time to bring about a change in the environment" (2006, p. 70).[7] Our case study in jazz is clearly a paradigmatic case of joint action on this definition. Likewise, Sebanz et al. (2006) suggest that joint action is supported by the ability to share representations though joint attention, predict each other's actions, and integrate the predicted effects from both one's own actions as well as the actions of the others taking part in the joint task. A variety of other mechanisms have been proposed for joint action as well, from those realized at the level of embodied sensorimotor coupling up to level of social conventions and cultural transmission (Vesper et al., 2017).

A full implementation story for shared P-intentions, especially in the case of ensemble musical performance, requires additional considerations about the existence of shared goals. As noted by Elisabeth Pacherie and Jérôme Dokic (2006), for instance, during jazz performance "the participating agents *share the same goal*, creating a new version of a jazz standard; they act towards that shared goal by playing their parts and they cooperate by adjusting their individual performance to the performance of other members of the band to achieve the common goal" (2006, p. 107, emphasis added). Peter Keller (2008) similarly argues that a coherent ensemble

[7] It is important to note that Sebanz et al. (2006) do not specify that a shared intention has to be in place in order to have a joint action. I grant this point for a variety of basic kinds of action. However, considering the nature of jazz music and a group navigating a song together, I suggest that at least jazz will require some sorts of shared intentions to explain its particular features.

performance is supported by a combination of anticipatory auditory imagery, prioritized integrative attention (i.e. the combination of monitoring other and self outputs during ensemble performance), and adaptive timing. For Keller, all of these mechanisms are aimed at promoting "ensemble cohesion" in light of a common goal. I shall return again to consider shared and common goals in more detail below.

Moving to the level of the motor plans and processes, the role of shared M-intentions is concerned with motor resonance, which is an important part of how we predict the actions of others (Gallese, 2005). Motor resonance, which has also been called motor simulation, refers to the ways in which a person's motor system activates both when performing an action and when watching another person perform an action. There has been evidence that the activation of the motor system is sensitive to intentions and goals, even in rather abstract contexts (Becchio, Manera, Sartori, Cavallo, & Castiello, 2012; Iacoboni et al., 2005). Additional evidence of the specific mechanism comes from research on the Mirror Neuron System (MNS) (Rizzolatti, Fadiga, Gallese, & Fogassi, 1996).

A recent study by Laura Schmitz, Cordula Vesper, Natalie Sebanz, and Günther Knoblich (2018) provides further evidence that motor resonance occurs when people do, as well as do not, consciously attempt to coordinate specific features of their actions. They found that resonance may even occur in times when it may be detrimental or irrelevant to the task at hand. To test this and related hypotheses, the authors measured the interference effect from co-representing the aim of a conspecific's action sequence, and in a context where the features being tracked are seemingly irrelevant for the completion of the joint action at hand. Moreover, Schmitz et al. (2018) provide evidence that co-representations track the order of temporality in action sequences.

The layout of the first experiment, which was utilized for the subsequent two experimental variations and slightly altered for the last three, had two people seated side-by-side behind a table. Each person was given a wooden dowel and faced a column of four 5cm markers placed in front of them, with the space between each marker at an equidistant 13.3 cm. In some experimental conditions, agents were asked to move along a similar path from the first to the fourth marker. In these cases, both person A and person B would be tasked to reach the final marker simultaneously while stopping at the third marker along the way, for instance. In other experimental conditions, the final task of simultaneously reaching the fourth marker remained the same, however, person A was required to go to the third marker along the way, while person B was required to go to the second marker. Later variations of the experiments were similar but controlled for various confounding factors, such as whether certain features of the markers, including size and colors, played an essential role in how they were represented, as well as whether or not visual contact with the other participant makes a difference.

Results across the different experiments found that co-actors consistently moved slower if the same action was performed in a different order, in contrast to where the same action was performed in the same order. The authors also suggest one particular musical application in the case of choirs singing a round (2018, p. 78), since each member of choir is performing the same action at a different starting point.

The use of shared P-intentions and M-intentions may help address some of these results in practice. As action coordination and prediction are integral to both P- and M-intentions, this capability would make both important candidates as “coordination smoothers”. As noted by Michael (2017), coordination smoothers include any processes that helps simplify and streamline an interaction process. Some of these may occur at the level of P-intentions, such as the use of

gestures or stock musical phrases to make sure that the flow of the improvisation is maintained. Having shared M-intentions furthermore makes it easier to predict what other people might do next, which in turn will make the improvisation run smoother than it would otherwise.

Finally, as with the case of individual intentions on the DPM, the three levels of the shared DPM model are also dynamically coupled to each other (Tollefsen, 2014, p. 22). As a result, various features of higher-level intentions have downstream effects, and lower-level intentions feedback into higher levels during the unfolding of an action. The interrelation of these levels, in turn, is important concerning how action goals may be a part of M-intentions that are otherwise inaccessible to conscious access and awareness.

Is the Shared DPM Model Minimal Enough?

Thus far I introduced the DPM model with certain assumptions that representations and goals are an important part of the overall explanatory story. I have not yet explored if an even more minimalist option can operate without appealing to these conceptual tools. A critique of the representational DPM approach following this latter consideration will manifest in two ways. First, in a general form, enactivists have argued that we should replace representations with a dynamical, non-representational approach to understanding the role of intentions and intentions-in-action (Gallagher, 2008; Hutto & Myin, 2017). Second, as a concern with the specifics of music performance, according to Andrea Schiavio and Simon Høffding (2015; see also Høffding, 2019), we have good phenomenological evidence in favor of rejecting shared goals and shared representations for guiding collective music-making. I shall address these concerns in order below.

Shaun Gallagher raises several worries about the use of representations in cognitive explanations, although I shall focus here on the claim that “In the case of action[, the concept of

representation] is nothing more than a handy, but often confused and misleading term, a bad piece of heuristics, an awkward place-holder for an explanation that still needs to be given in dynamical terms of an embodied, environmentally embedded, and enactive model" (2017, p. 106). While I grant that many representationalist accounts have fallen prey to this worry and ended up being inert placeholders instead of offering the detailed explanations required to understand the cognitive phenomenon in question, I am not convinced that the use of representations in the shared DPM model can be summed up as "a bad piece of heuristics," for reasons I shall now introduce.

I take the most important issue on representation in the shared DPM model to be found at the level of motor representations. This level is of particular concern for two reasons. First, at the level of D-intentions, and perhaps P-intentions, even radical enactivists may accept that some sorts of representations are at play, at least insofar as cultural and linguistic components are used in creating and maintaining the intentional activity of music-making. Second, since M-intentions are the level where a dynamical explanation is an essential part of the cognitive story, establishing that representations are present here will be an important part of answering Gallagher's challenge for the usefulness of representations.

In explaining motor intentionality, Pacherie (2018) argues in favor of utilizing representations, in part following Jose Bermùdez's criteria for what counts as a representation: correctness conditions, compositionality, cognitive integration, and explanatory usefulness. Motor representations have correctness conditions insofar as they guide an unfolding action, and such actions can succeed or fail. There is a role for compositionality in the fact that different motoric actions chain together to achieve a goal. This compositionality is further shaped by the overall goals and intentions at the D- and P- levels and, thus, cognitive integration occurs as

well. Finally, motor representations provide a level of explanatory usefulness over and above a purely mechanistic explanation of action by explaining how and why different motor plans chain together.

While more could be said on each of these points, I want to focus more on the final criterion of explanatory usefulness. Pacherie specifically argues that such usefulness can best be seen when "the motor behavior they are meant to explain *could not be explained in terms of a lawful correlation between sensory stimulus and behavioral response*" (2018, p. 381, emphasis added).[8] The reason why lawful correlations cannot work here, according to Pacherie, is that the various intentions of the agent will have a direct impact on the way they engage with any environmental stimulus. They have this impact, moreover, in a fashion that is not reducible to a lawlike relation where X sensory stimulus will result in Y behavioral response. A jazz drummer who complements a soloist's previous phrase will act differently than one who has decided to push them by being a source of disruption instead. This result holds even though the drummer is responding to the same stimulus across both cases. Likewise, there is nothing intrinsic to the stimulus itself that allows us to definitively say which possible responses will, let alone ought, to be selected at any given time.

We can further supplement Pacherie's account of motor-intentionality with Maya Gratier's (2008) account of grounding in jazz performance. Grounding concerns the creation of a common ground, which is constituted by mutual knowledge, beliefs, and assumptions.

[8] Ecological psychology often takes an essential part of explaining perception in terms of lawlike explanations. I therefore return to consider this claim indirectly in the following chapter, especially in regards to the idea that the brain's primary purpose is to resonate with ecological information.

According to Gratier, grounding in the case of jazz is similar to other sorts of communicative contexts.[9] The similarity may even reach down to the level of basic caregiver-infant interactions.

In order to establish how grounding works in the case of jazz performance, Gratier analyzed several different sound plots from live performance to study "the musicians temporal contouring of expressive intentions" (2008, p. 89). The analyses included independent sound-only and visual-only analysis of the interaction between a drummer and a guitarist, and established various indications of musical grounding that occurred during the performance. Of particular interest is the ways in which the drummer and guitarist studied by Gratier followed each other with patterns of embodied activity, especially insofar as they embodied acts of performance in particular ways (2008, p. 92-4).

These embodied activities follow important ebbs and flows from tension and release throughout the performance (a topic which we will return to again in chapter 4 on Predictive Processing). They are likewise good candidates for shared M-intentions that necessarily represent features of the performance and the actions of others while a musician is deciding what to do next. In a similar vein, several studies by Ashley Walton and her colleagues (Walton, Richardson, Langland-Hassan, & Chemero, 2015; Walton, Richardson, Langland-Hassan, Chemero, & Washburn, 2018) traced the thermodynamics between improvising duets and found important roles for the embodied interactions of performers navigating their collective and individual understanding of the piece being performed.

An enactivist may here push back that the previous two examples do not provide support

[9] It is interesting to note that the empirical evidence drawn in the third section of Gratier (2008) is focused on a duo performance between a drummer and a guitarist. While I do not argue that the duo focus results in an artifact of overlap between other communication contexts and jazz improvisation, I would stress that there may be an important shift in the sorts of processes involved once a duo becomes a trio, a quartet, a quintet, etc. Whether or not this intuition holds is an open empirical question.

for a representational account as proposed by Pacherie. While I do not think that Pacherie's argument is a knock-down reply to Gallagher's worries – indeed, Pacherie herself notes a similar point (2018, p. 384-5) – and the above is not meant to be an argument against the existence of alternative, non-representationalist explanations, I suggest it provides enough evidence to at least see how the DPM model can utilize representations in a coherent and explanatorily useful way. The explanatory usefulness of motor representations is further realized in conjunction with motor resonance as considered above. Finally, since the nature of P-intentions and M-intentions on the shared DPM model are necessarily dynamical in nature, they at least attempt to directly speak to Gallagher's call for offering a dynamical explanation of action.

Turning to shared goals, Schiavio and Høffding (2015) leverage a series of arguments against the need to posit collective goal states for successful musical performance. Focusing on the case of musical absorption in the Danish String Quartet (DSQ)[10], they argue that there is phenomenological evidence for players successfully performing without taking notice of the mental states of others nor developing a preexistent shared goal. More specifically, in their words, "ensemble performance can take place without attention to either shared goals, or to the other ensemble musicians" (2015, p. 4). Schiavio and Høffding ground their argument against three core claims about joint musical attention on representationalist accounts, namely the representationalist views that: (1) musicians playing together must have some awareness of their co-players, (2) this awareness must be focused on the subjective states of the other musicians during performance, and (3) there must be some common goal during performance among musicians (2015, p. 10). While they are not targeting the extended DPM model specifically, these three claims are in line with what the DPM account entails for an account of jazz

[10] Høffding (2019) claims that a similar move can be made in the case of jazz improvisation.

improvisation, and, as such, it will be important to reply to Schiavo and Høffding's concerns in order to defend my application of it. However, since Schiavio and Høffding ultimately grant that the first claim is correct to a modified degree, I will focus on the second and third points below.

For evidence against (2), Schiavio and Høffding point to two particular cases from their interviews. First is a situation where one DSQ violinist, Frederik, had been crying during performance while the other violinist, Rune, did not notice. Second, in a different performance, Rune experienced shoulder pain to a point that was noticeable to some of the other quartet members, yet Frederik didn't notice at all. As a result, Schiavio and Høffding conclude that these cases support the idea that at least the 1st violinist is able to perform without awareness of his co-player's mental state.[11] Moreover, they argue that "The example above demonstrates that awareness of others' mental states, or a process of 'putting oneself in the other person's shoes' *is neither necessary nor sufficient* to account for the kind of awareness required for JMA [Joint Musical Attention]" (2015, p. 11, emphasis added).

In contrast, I suggest that these examples do not yet establish the strong conclusion that Schiavio and Høffding wish to draw, especially since there may be other subjective mental states that can be targeted in addition to the emotions of co-performers. For instance, it is possible that the 1st violin player was continually tracking the intentions of their fellow performers without tracking their emotional state, and the phenomenological evidence provided does not disprove that claim. As a result, while I don't suggest that there can be affectless intentions, I do suggest that it is possible to track important parts of another person's mental happenings without tracking emotions and/or all of their mental states.

[11] Rune and Frederik alternate between playing 1st and 2nd violin; the moments of lapse in both cases came when they were playing 1st violin.

Schiavio and Høffding foresee my appeal to different mental states and consider two additional points against it. First, they hold that it is far from obvious what sorts of subjective states are required for performing a piece of music and taking the time to figure these out, they hold, will not be a useful exercise. In some important ways, this argument mirrors Gallagher's (2017) facetious economic argument against representations. The core of Gallagher's argument is that the amount of work it would take to, first, agree upon a specific notion of representation and, second, explain how this notion would function in a cognitive explanation is far too costly to pursue, especially with non-representationalist alternatives available and continuing to be developed.

While I agree that we shouldn't spend the majority of our efforts cashing out necessary and sufficient conditions around the kinds of subjective states that could be part of musical performance, I nevertheless believe we already have at least one clear candidate for a necessary subjective state to be tracked, namely the intentions of co-performers. Furthermore, while the tracking of these intentions may not rise to conscious awareness, as in the case of tracking M-intentions or developing shared M-intentions, I suggest that they are nevertheless essential features of performance. An additional upshot from this consideration is that, if it is possible to track intentions without being able to consciously access that one is doing so, it is similarly possible that the phenomenology during performance will not reflect what is nevertheless an important aspect of the cognitive capacities and processes utilized therein.

Second, Schiavio and Høffding argue that the awareness of subjective states would "surely…be affected by strong emotion such as crying" and furthermore "it is hard to conceive of a case in which crying would not affect the emotional quality of the music" (2015, p. 12). While these assumptions are intuitively plausible, they are little more than intuitions as currently

stated. Moreover, considering the level of skill and experience held by all members of the DSQ, it is conceivable that a musician may experience a strong emotion without conveying it in a way that comes across in the performance. They may also be able to bracket off the feeling of sensing emotions from others during performance to a certain extent as well. In an analogous case concerning direct perception of emotions, we don't say that the inability to directly perceive all emotions is a counter to the theory that emotions are sometimes directly perceive; likewise, the idea that the DSQ members attend to the subjective states of their co-performers is not defeated if we find counter-cases where (certain) subjective states are not perceived during performance.

Turning from tracking mental states to the notion of a collective goal, Schiavio and Høffding define goal by quoting Keller's (2008) definition in full that:

> once performance goals are established, they reside in memory as idealized mental representations of the sounds constituting the musical piece. Performance goals embody a performer's intentions and expectations about how his or her own sound *and the overall ensemble sound* should be shaped dynamically over time. With such goals in mind, musicians develop performance plans (usually during private practice) that guide the motor processes involved in translating the goal representations into appropriate body movements. (p. 206, emphasis in original).

A key element of this specific definition is that the goals are static instead of dynamic, especially as they are tied into performance plans developed ahead of time instead of in the moment of creating the song. For evidence against the shared goals of (3) above, Schiavio and Høffding refer to an interview with the DSQ cellist Frederik about why he played a certain phrase in a particular way. According to Frederik, he is unable to predict how any of his choices about music will play out ahead of time. He further refers to this feeling along the lines of surprise at the

outcome, which, Schiavio and Høffding submit, would be an unexpected result if he were working with performance plans and goals as defined by Keller. Moreover, they note that (Schiavio & Høffding, 2015):

> the goal, being instantiated by the dynamical interplay between performers and music, cannot be reduced to the idea of the performance stored in the memory of the musicians, or to a unified concept of sound – rather, it must be constantly open to various interpretations and modifications that stem from the sensorimotor loops from which it is dynamically and intentionally constituted. (p. 14).

In turn, Schiavio and Høffding suggest that the interview evidence from Frederik counts against positing either shared or individual goals. Since there are not goals informing the individual musician, there will not be even higher-level goals informing the collective.

I agree that the notion of goal provided by Keller (2008) is problematic, especially in the context of jazz improvisation. An essential aesthetic feature of bebop is in part not agreeing upon some sort of robust, explicitly collective goal of how the song will unfold at the outset, after all. Yet, at the same time, it doesn't seem that the idea of a goal in itself has been refuted, but rather a particular sort of collective goal. Furthermore, even Schiavio and Høffding allow that there may be a shifting sort of goal that underlies performance.

An additional issue against Schiavio and Høffding concerns the fact that the relevant co-representation of goal states may not be present in conscious experience, as we saw in the case of the Schmitz et al. (2018) study above. As a result, in order to establish that there is in fact no collective or shared goals, overlapping goal states, or simulations of the mental states of others at play in performance, more evidence needs to gathered to supplement phenomenological considerations. The possibility of tracking the music itself and the mental states of others without

conscious awareness will also have import for our considerations to what is happening in cases of musical absorption, as considered in several areas above. In these cases, while a performer may not be consciously aware of the fact that they are attending to the mental states of others, they nevertheless do so.

In a similar manner, and to return to the language of the collective DPM model, the ensemble may still be guided by a preset, general D-intention(s) even as the specifics of that intention change during the act of performance. Morphing the content for various intentions across the DPM levels during group performance does not establish that there are no preexistent intentions, just as it does not establish that there is no attendant planning or goal states during performance. It only establishes that intentions and plans can change, sometimes to a great degree, during the actual moment of performance, in contrast to what was agreed upon beforehand.

My main point from the previous two paragraphs can be turned on its head, however. In this reformulation, my appeal to subpersonal goals is problematic because the notion of goals simply cannot be applied at that explanatory level (Høffding, 2019, Chap 11). In Høffding's words, the notion of goals furthermore "indicates a unidirectional causation of musician-causing-music and overlooks the performatively passive sense in which the music plays itself and 'becomes its own ground' irrespective of one's goals" (2019, p. 227).

There are at least two avenues available to reply to these twin worries. First, considering the intention cascade on the DPM model, I grant that the notion of goal may not appear in isolation at the lower-level of intentions or in a completely unconscious process. It instead becomes relevant insofar as lower levels of the intention hierarchy are simultaneously constrained and inform the goals of the person as a whole, even at the personal and conscious

levels of explanation, respectively. Second, it is possible for a person to have a directedness to particular goals that are nevertheless open to multidirectional causality, as we saw in chapter 1 concerning the definition of improvisation and an open-ended navigation of the environment.

There is also additional relevant empirical evidence for how performers track and co-represent the mental states of their fellow performers' goals during performance. EEG evidence from several studies (Kourtis, Knoblich, Woźniak, & Sebanz, 2014; Loehr, Kourtis, Vesper, Sebanz, & Knoblich, 2013; Schmitz et al., 2018) offer evidence that people often (subconsciously) make predictions about their co-actors' actions. In addition, a study on how expert musicians can predict suboptimal timing of a novice during a duet by Thomas Wolf, Natalie Sebanz, and Günther Knoblich (2018) found that experts utilized both knowledge of the score and knowledge of their co-performers performance to help navigate the coordination task.

Furthermore, in the Wolf, Sebanz, and Knoblich (2018) study, the authors suggest that there are two main factors needed to facilitate adaptation of the experts to the suboptimal timing of novices. On one hand, in easier passages, experts are able to defer to familiarity with novices' performances and adjust their own timing accordingly. On the other hand, in more difficult passages, experts may defer to familiarity with the scores available to the novices. The impact of these two factors was tested by a running a series of eight blocks during the experiment. Each block consisted of novices performing their part of the duet eight times. For the first four repetitions, experts were either able to hear the novices' performance or not and either did or did not have access to the sheet music for the novices' part. For the second four repetitions, the experts were asked to match the keystrokes of the novice as closely as possible.

Wolf et al. (2018) further suggests two main mechanisms used by experts for addressing the suboptimal timing of novices. The first is that experts use their experience with a novice's

performances to develop an error matrix that augments their predictions for timing. The second is to offload the task onto the symbolic information encoded in the score. While other mechanisms may be relevant for these situations, this study provides some information as to how musicians with different skill levels can nevertheless track and respond to the intentions of co-performers.

Augmenting the DPM Model with Interkinesthetic Affectivity

In the previous section, I considered the enactivist account provided by Schiavio and Høffding as a critique of the extended DPM model. However, unlike the maximalist accounts of Gilbert and Bratman canvased earlier, both the DPM model and enactivist alternative operate as minimalist accounts. As a result, it's possible that they may prove more harmonious than not, even if they are not fully reconcilable in their current forms. As a result, in this section I shall consider the possibility of further extending the DPM model with interkinesthetic affectivity, a concept developed in Høffding's book on the phenomenology of musical absorption (2019).

In an explanation of joint action and collective cognition, Høffding suggests three different mechanisms for how the DSQ performs together (2019, p. 229-230). Two of those, motor-resonance and explicit coordination, have been considered in regards to the extended DPM model above. As a result, I will focus on the third mechanism – "interkinesthetic affectivity" – and see if it acts as a mechanism that cannot be captured in DPM terms already specified.

Interkinesthetic affectivity is knowledge developed through a specific feeling. This feeling, as a kind of we-agency, is deeply rooted in the trust that is developed as a group of musicians plays together over an extended period of time. It also results in the feeling of "one shared body," through a combination of auditory perception, affectivity, interoception, and

intersubjective proprioception (Høffding, 2019, p. 235). This final ingredient of intersubjective proprioception is moreover taken to be an interkinesthesia, or a joining together of body schemas. The DSQ is an excellent example of a kind of group that can experience interkinesthetic affectivity, having spent tens of thousands of hours practicing and performing together over the course of many years.

The full analysis provided of these four features underlying interkinesthetic affectivity is a lucid, engaging, and useful tool for understanding music cognition. Moving through it in detail will unfortunately require more space than currently available. However, one particularly important point about this addition is that it offers an alternative to some of the representation-laden language that has been used in developing the DPM model. The two things that it adds on this front are highlighting more strongly a role for affect and considering the joining together of body schemas, in contrast to merely tracking the bodily actions and intentions of others. Whether or not one or the other of these explanations – i.e. representation heavy or pure interkinesthetic affectivity - will prove more explanatorily fruitful is an open question. Likewise, whether or not one needs to give up on all notions of representation to talk about interkinesthetic affectivity is an additional open question for the future.

It is worth noting in conclusion that the picture presented by Høffding mirrors important parts of the phenomenon described on Gilbert's account. This relationship is even more clear if we put it side-by-side with Hagberg's use of the plural subject to analyze the Stan Getz performance. While there are important differences to Gilbert's account, using interkinesthetic affectivity to understand novice jazz performance may still fall prey to the same worry that we have successful novice performances without such interkinesthetic affectivity. As a result, it has

a more limited application to contexts of expert performance, rather than performance in all different levels of skill.

Chapter Conclusion

In this chapter, I considered whether there is a distinct category of shared intentions that underlie jazz performance. After providing reasons for thinking that such shared intentions exist and are irreducible to individual intentions, I considered three models of shared intentions. These three models could further be grouped into two maximalist accounts and one minimalist account.

While each of the models offered important insights for different features of bebop jazz performance, such as the nature of highly skilled and meshed performances from Gilbert's account and the role of planning and negotiations for structuring a performance from Bratman's account, I ultimately suggested that the extended DPM model provides the best model for shared intentions in bebop performance regardless of the skill levels of those who are taking part. I also considered concerns that the extended DPM model is ultimately not minimal enough, responding along the way that musicians track the mental states of others in a nontrivial number of cases and often have a shared goal at hand during performance, even in an improvised performance.

One standing assumption of this chapter was that representations are a relevant conceptual tool for our psychological explanation because lawlike explanations are not suitable alternatives to describe how perception and action work during bebop improvisation. In contrast, one branch of psychology that rests heavily on the idea that perception and other psychological processes are lawlike is ecological psychology. Considering the important alternatives available in this area, it is time to move past this standing assumption and see how an account of jazz improvisation can be developed in ecological terms. I will therefore turn to this new avenue as an alternative way of capturing an essential feature of bebop performance in the following chapter.

Chapter Three: Jazz, Affordances, and an Ecological Approach to Improvisation

Drawing from research in ecological psychology, the concept of affordances has been utilized to explain aspects of both music perception and production (DeNora, 2000; Clarke, 2005; Linson & Clarke, 2017; Windsor & de Bélzac, 2012; Burland & Windsor, 2014; Krueger, 2014; Menin & Schiavio, 2012). In this chapter, I consider how ecological approaches have been applied to understanding musical performance, especially of jazz improvisation, and offer further considerations about the nature of affordances, ecological information, and resonance to help develop ecological psychology in general, along with specific accounts of music performance.

The concept of affordances was originally introduced by James J. Gibson. According to Gibson (1979):

> The affordances of the environment are what it *offers* the animal, what it *provides* or *furnishes*, either for good or ill. The verb *to afford* is found in the dictionary, but the noun *affordance* is not. I have made it up. I mean by it something that refers to both the environment and the animal in a way that no existing term does. (p. 119, emphasis in original).

Based on this particular definition, affordances apply to a wide set of relations. As such, they are not limited to cases of perception in basic contexts, such as seeing that a slope affords climbing or a chair affords sitting, but expand to all aspects of an organism's niche, including patterned sociocultural practices such as running a trail race or music-making. In addition, musical affordances further encompass the fact that music affords emotional control and manipulation of affect for a listener (DeNora, 2000; Krueger, 2014) along with the more abstract fact that musical motifs have come to symbolize certain character traits and, as a result, film music affords an audience the ability to make particular judgments about a character.

To further explain how these different types of affordances refer to both environment and organism, consider the case of a standard metal folding chair. To an able-bodied adult human, this object affords sitting, at least in the right situations with the chair open and on stable ground. The same affordance categorically does not exist for ants or elephants. The chair may afford throwing in many cases for humans and elephants, yet again not in the case of ants. Finally, however, if the chair were made of wood and the ants in question were termites, it would afford eating to ants but not to humans or elephants.

Moving from affordances for action to why we choose to act in certain ways, we find an important ecological account of agency that has been developed by Mark Risjord (2014). In what follows, I will use Risjord's account to both analyze a case study of Coleman Hawkins' famous 1939 recording of *Body and Soul* and to introduce several key features of ecological psychology that are relevant to explaining jazz improvisation. For his account, Risjord draws on the overlap of attunement, affordances, and meta-cognition, all of which will be defined and explored in the following sections.

As a result, after beginning with an introduction to various historical and musical details of Hawkins' performance, I introduce Risjord's account as it currently stands. The remainder of this chapter then focuses on a more detailed explication of each of the three core concepts from Risjord and related philosophical issues around them. For attunement, the notion of ecological information is considered in more detail, especially around the issue of ecological invariants and lawlike information; in regards to affordances, the different ontological conceptualizations of affordances – dispositional or relational - are introduced and the relational account is defended as the preferable alternative; finally, for meta-cognition, the idea that the brain resonates with environmental information is critically introduced and considered in detail.

A Case Study in Body and Soul

The jazz standard *Body and Soul*, originally composed in 1930 by Johnny Green with lyrics by Edward Heyman, Robert Sour, and Frank Eyton, has been dubbed by Ted Gioia as "the granddaddy of jazz ballads, the quintessential torch song, and the ultimate measuring rod for tenor sax players of all generations" (2012, p. 46). That the song would earn this distinction was not apparent when it was first composed. While there were several recorded versions from the early 1930s, most notably Louie Armstrong's take on the standard, *Body and Soul* only came into its own a decade later and flourished further as a bebop standard of choice, and beyond.

While many of these bebop recordings provide fertile ground for further analysis, my focus will be on Coleman Hawkins' 1939 version of *Body and Soul*.[1] Hawkins' solo on this track is considered a watershed moment in jazz and one of the great improvised solos in the history of recorded music. It was especially important in the development of jazz between the big band and swing jazz era of the 1930s into the bebop era of the 1940s. In addition, while not yet working within a bebop idiom, certain parts of the performance (especially the two-chorus improvisation by Hawkins that lasted the majority of the take and mostly hinted at the melody without explicitly stating it) have led critics to see it as a proto-bebop style.

Hawkins begins his solo after a short piano introduction. One immediately striking fact is that, after this piano intro, little of the song's original melody remains. Almost the entirety of the song instead becomes an extended tenor sax solo, lasting 2 minutes and 48 seconds on a track that ends at just over 3 minutes total. There are several additional reasons why this track is unique. To see some of them more clearly, we can compare it to Louie Armstrong's 1930 version of *Body and Soul*. While Armstrong similarly begins with a lengthy trumpet solo that lasts 50

[1] For an example of this version, see https://www.youtube.com/watch?v=zUFg6HvljDE

seconds, it segues into an instrumental chorus, followed by a verse as sung by Armstrong, then another brief solo and the outro.[2] The accompaniment behind Armstrong's and Hawkins' respective solos is also different. While the supporting music in Armstrong's version often follows the melody, Hawkins' accompaniment is a sparser affair, mostly consisting of accented beats and a straightforward drum pattern, with an occasional set of long notes from the horn section at various points.

There are multiple ways to run additional analyses on Coleman Hawkins' performance, from a focus on music-theoretical aspects to foregrounding the socio-cultural and historical factors that resulted in it. For my current analysis, I will turn to conceptual tools gathered from a qualitative study of improvising musicians conducted by Karen Burland and Like Windsor (2014), which are related to ecological psychology. I will also focus on what Hawkins' is doing during his solo without detailed reference to the rest of the group. While we shall return to the group aspects of bebop later in the chapter, I focus the aspects of an individual performer first to give a sense of how some crucial ecological concepts can be applied.

Burland and Windsor highlight an essential combination of two different aspects – behavior settings and affordances – that come together within improvised ensemble performances to make them possible. Behavior settings can be broken into three groups: *topographical features*, which include the layout of the ensemble during performance and the layout of the performance space itself; *climatological properties*, such as the temperature of the performance space; and *sociocultural practices*, including conventions such as how passively or actively an audience should respond when attending a performance. All of these settings have a direct impact on the ways in which music is made. In the words of Burland and Windsor, "the

[2] For an example of this recording, see https://www.youtube.com/watch?v=YpGZBr-RYK8

behaviour of musicians is afforded by relationships with their behaviour settings, and it is important for a psychology of performance to investigate these settings, the information they furnish and the use we make of them in performance" (2014, p. 103).

We can further supplement this list for affordances in music performance, following Mark Reybrouck (2012, p. 405), by taking consideration of musical instruments, playing techniques, and modulatory techniques, which are respectively related to the nature of the material used in the production of the instruments, the production of musical sounds, and the shaping of those sounds during performance. We can likewise consider the possibility that other musicians and the audience are also affordances during performance, although there is here an important open question of just how to understand the relation between all of these different kinds of affordances.

If one wished to run a full analysis of *Body and Soul* following in Burland and Windsor's and Reybrouck's footsteps, they would accordingly need to give a full account of all these features. For my current purposes, I will skip an analysis of the topographical features, which would entail exploring the exact layout of the performance space in which Hawkins' was situated while recording, as well as climatological properties, such as the exact temperature and humidity of the studio on October 11th 1939. Instead my focus will be on affordances and sociocultural practices.

There are multiple affordances available in Hawkins' environment while he was recording. Chief among them for performance purposes are his saxophone and other musicians. Some affordances of the sax, such as the ability to throw his instrument or use it to play a bass drum, are of little relevance for analyzing this particular performance. Others, such as the fact

that a saxophone can be blown to produce certain tones and these tones can be modulated by various finger patterns and changes to air flow, are of direct relevance.

These directly relevant affordances are further based on Hawkins' skills with his instrument as developed over years of study, play, and practice. They are tied into more particular features of a given performance as well, such as how tired Hawkins' was while making a recording or if the heat of the stage made him sweat to an excessive degree. Across all of these cases, moreover, Hawkins' standing as an expert allows him to focus on different affordances without paying as much attention to other parts of performance, such as specific fingering patterns or how to hold the horn. The saxophone, in other words, is incorporated into his body schema and treated more like an extension of his body than something outside of himself (Høffding, 2019). Yet there will also be cases where incorporation breaks down for various reasons, even for an expert, and especially when we reintroduce climatological and topographical considerations to the analysis of a given performance.

It is also important that we are not just considering saxophones in general, but a tenor saxophone in particular, and Coleman Hawkins' tenor sax to be even more precise. Some of the affordances he engages with are related to the intricacies of this particular instrument, along with various recording technology used to capture the performance. At the same time, becoming an expert also makes it easier to find similar affordances even while playing a stranger's instrument.

We find additional changes when we look at the sociocultural settings that link Hawkins to his instrument and the musical material he was playing. When Hawkins' recorded this version of *Body and Soul*, he had just returned to America after spending several years in Europe. One thing he noticed about the music scene in NYC and US at the time was that little had changed

between the licks and musical selections being used in tunes when we had left in 1934 and upon his return in 1939.

The musical ideas that he put forward with this performance were not met with immediate acclaim from all of his peers and jazz fans, however. In Hawkins' words (as cited in *National Public Radio*):

> I thought by the time I came back, I thought the musicians here would be much further advanced, But they were just like when I left. I didn't see any — nothing. They hadn't advanced; they hadn't done anything. A lot of them used to say I was playing wrong notes. Like the first time I played 'Body and Soul,' when the record first came out, well, everybody said I was playing the wrong notes in it. It was funny to me. They just weren't making these changes in 'Body and Soul.' It's the only changes to make in 'Body and Soul.' That's what... I couldn't understand why wasn't they making them. (2000).

There are at least two main points that can be distilled from this quote. First, in regards to "these changes in 'Body and Soul'," we find that Hawkins' focused his interpretation of the song along the lines of chord changes, a practice that would shortly thereafter become the bedrock for bebop performance.

Second, we find an important shift in how Hawkins' relates to the musical material that is distinct from many of his predecessors and contemporaries. Embedded in this quote it is clear that Hawkins' does not identify *Body and Soul* in terms of a melody or lyrics. Instead, he takes its foundation as the chord changes, which in turn form the basis of his improvisation.[3] A related point can be considered regarding other people at the time thinking he was playing "wrong

[3] This point can also be captured as shifts among different types of improvisation introduced in the Brown et al. (2018) taxonomy discussed in Chapter 1.

notes." Similar things were often said about other proto-bebop and early bebop players, especially in the case of pianist Thelonious Monk, and they highlight how sociocultural settings must consider both what is afforded by a particular sociocultural milieu as well as what changes must happen to make different practices possible, which has ramifications for the possibilities for relating to musical material as well.[4]

An Ecological Model of Jazz Performance

Thus far we have considered *Body and Soul* as a case study, along with some conceptual tools from ecological psychology for analyzing it. I now turn to expand this analysis further by introducing a specific model that captures improvisation developed in the work of Mark Risjord. Risjord's (2014) account of improvisation combines three key features - *affordances*, *attunement*, and *meta-cognition* - all considered under an *ecological theory of agency*.

In line with quick-paced activities, such as fast breaks in basketball or catching fly balls in baseball, Risjord notes that jazz improvisation and improvisation in general are best understood along an ecological theory of agency. In contrast to situations where performers are taken to merely reproduce preexistent structures, jazz improvisers create something new as well as respond to the current moment. Such responses in turn rely on the fact that musicians, especially professional musicians, are skilled and situated in their particular performance setting.

Instead of suggesting that the agent has to think ahead in detail for acting on the environment, Risjord proposes that the processes underlying these situated responses are realized

[4] Before turning to Risjord's general ecological model of improvisation, a few words on the role of recording vis-à-vis improvisation is important (Brown, 1996). Recording technology allows for an improvisation to be studied over and over. On one hand, this result has pedagogical benefits since it becomes possible to focus on various techniques that musicians use to navigate songs; a budding musician can compare and contrast Hawkins, Armstrong, Fitzgerald, and countless others in their renditions of *Body and Soul* while learning to play. Learning these licks and tunes will in turn be important for a novice as they develop their own unique voice. On the other hand, recordings highlight a tension about the nature of improvisation, since the moment that you listen to the same improvisation again, something important about it has changed.

in action when the performer attunes themselves to the environment. Such attunement is furthermore aimed at affordances that are relevant for the individual, as well as those that are relevant for different members of the group. Meta-cognition comes into play when musicians track their own musical output, along with the output of the other musicians, and attempt to guide their performance in relation to an overall group sound.

Before returning to these key three features again, however, more must be said about what Risjord means by an ecological approach to agency. It is often assumed that we should begin with an account of action in order to draw conclusions about agency. This sort of presupposition is at least partially operant in Chapter 2 above, where I started with the idea that jazz improvisation is a kind of action, then used intentions to explain exactly what kind of action it is. In such cases, intentions make actions distinct from mere happenings; the exploration of what intentions moreover are can tell us something important about action and, in turn, agency.

In contrast, by taking agency as fundamental, Risjord's account of agency moves from how we understand agency to how we are able to act. In doing so, Risjord considers several accounts of agency from a core debate in the social sciences about how individuals relate with social structures. The main tension in that debate rests on the difference between microfoundational approaches, on one hand, and practice theory approaches, on the other.

Microfoundationalists understand social structures to be the product of deliberate individual actions and choices, whether or not the structures in question were built intentionally or unintentionally. Agency for microfoundational theorists, in turn, takes individuals as "choosers" that follow the dictates of practical reason. A specific example here is how rational choice theorists consider the nature of agency in prisoners dilemmas, namely as the product of individuals making the best choices to serve their own personal interests.

Practice theorists, including Anthony Gidden and Pierre Bourdieu, reject the microfoundationalist understanding of both social structures and agency. Instead, social structures are best understood as "recurring patterns of interaction" and regularities. Having agency on this account is thus about how an individual navigates within reproduced patterns and regularities, along with building up some sorts of differences along the way.

The ecological account of agency takes pieces from both microfoundational and practice theory approaches. From the former, it takes the idea that social structures are continually remade and not mere reproductions of regularities. This point is especially salient for our understanding of jazz improvisation, since a main issue is how improvisers make novel material, or at least create different material on the spot, instead of merely reproducing parts of the standard being performed or entire licks from memory. From the latter, it takes the idea that we should reject the view of agents as deliberate, rational choosers. Furthermore, instead of being focused on the internal processes of an individual, agency is better understood in light of mutual responsiveness and coordination between individuals, plus an individual attuning to features of others and social structures when they act.

Ecological agency entails flexible engagement with the environment. This flexibility, in turn, requires meta-cognition to guide different ways of engaging with it. In Risjord's words, meta-cognition "is our capacity to bring to mind prior plans, explicit beliefs about the environment, knowledge of explicit rules, and interpretations" (2014, p. 234). Some parts of jazz performance may proceed without meta-cognition in an important role. Yet there remains an important role for it, since some musicians report a conscious stance towards parts of their solos and insofar as attunement can be augmented with personal choices on the matter. This latter point can be seen in how a performer plays radically different solos for the same piece over time.

Hawkins, for instance, played *Body and Soul* a multitude of times after his 1939 recording, sometimes quoting various parts of this classic performance and using it as a springboard for additional improvisation.

It's further important to foreground that attunement is not necessarily focused on creating harmonious outputs. For instance, in exploring how creativity is a distributed process during jazz performance, Adam Linson and Eric Clarke (2017) argue that any model of improvisation must move beyond assumptions solely or primarily grounded on convergence among different performers. While such convergence is an important part of the process, there will also be divergent processes going on as well. One of these divergences is the fact that individuals are focused on their own performances and selectively focused to other parts of the performance. In their words, a narrow focus on mutual collaboration "underplays the simultaneous listening-while-performing that takes place *in parallel* among players, rather than convergently" (2017, p. 59, emphasis in original). Such parallel processing allows players to afford distinct possibilities to others based on the choices that they make during performance, especially since those choices shape the overall contours of the music in the moment.

The remainder of this chapter will focus on developing additional theoretical issues around attunement, affordances, and metacognition, respectively. More specifically, for attunement, we will focus more on its nature and scope of ecological information that is the target of attunement in the first place. For affordances, we will adjudicate between conflicting metaphysical accounts of whether they are better understood as dispositional properties of the environment or relational features of the organism-environment system. For metacognition, we will consider whether ecological psychology has done enough to capture what, if any, internal processing is necessary in cases of ambiguous or underdetermined environments, with a

particular focus on the metaphor of the brain as a resonance organ being considered in detail. Prior to these analyses, however, I will take a brief detour through an even more general account of ecological acoustics to set the stage for these considerations.

Ecological Acoustics

In *The Senses Considered as Perceptual Systems*, Gibson (1966) discusses several important features of the auditory system, including the inner ear, outer ear, and the cochlea. He also stresses that auditory localization is possible insofar as we have ears spaced on both sides of our head. The spatial difference between ears means that sounds hit them at different times, and through that information an organism can orient themselves to their origins. The sounds will further be augmented by how a person engages with objects in the environment, as well as the medium in which soundwaves are moving. We will return to these points later in this section.

The act of articulating sounds is integrally tied into discriminating sounds. A similar point could be made in the case of music as well. In Gibson's words (1966):

> The young child learns to guide the flow of speech sounds by auditory feedback. The adult can learn to guide a flow of instrumental sounds in much the same way. In either case the pattern and sequence of muscular action is controlled for a purpose. The blowing of flutes, the 'lipping' of horns, the plucking, bowing or fingering or strings, and the flipping of drumsticks are not, in principle, unlike the articulation of the voice. (p. 96).

While studies on proprioceptive feedback and musical performance have found important contrasts between speech and music-making, regarding differences between the two phenomenon and how the link between auditory perception and proprioception operates for playing different instruments (Pfordresher & Palmer, 2002; Pfordresher, 2008), what remains

throughout is an important connection between various channels of feedback in order to maintain a stable performance across time.

One detailed account of ecological acoustics comes from the work of William Gaver. In a series of articles (1993a; 1993b), Gaver explored what we hear and how we hear in ecological terms. The what side of the equation focused on charting the sounds that different materials will make across mediums and contexts. The how side of the equation considers how an organism can move from proximal sounds to distal causes in the environment. For both what and how, Gaver stresses that we hear an auditory event. Audition is not just distilling a specific percept among a stew of sounds; rather, it is about perceiving an overall event in which all kinds of listening takes place.

We can explicate Gaver's account through his example of hearing a car approach (1993a, p. 6-9). Assume you are standing by a road and hear a car driving in the distance. In this case, the car itself is the source of a sound-producing event. This event, in the words of Gaver, "provides information about an *interaction* of *materials* at a *location* in an *environment*" (1993a, p. 6, emphasis in original). The various components of the car are moving together in a way that moreover provides a *meaningfully* structured sound. We may not be able to isolate the individual pistons, axels, wheels, and belts of the car whirling and twisting away, but we can hear sounds as produced by the interaction among these and other car parts. This intermingling creates sound waves that bounce off parts of the environment as they travel from the approaching car to your ears. Bouncing about in the environment helps structure and shape the sound waves that reach us. The combination of sounds plus their variable intensities and pressures develops an auditory array, which allows for localization of sound and distinguishing the kinds of objects in the environment.

Gaver is clear that his account is focused on everyday listening in contrast to musical listening. In traditional understandings of everyday listening, we listen to the source of a sound. The car barreling down the road is important because it's a car and I'm on or near the road. In musical listening, we focus on particular qualities of the sound source. Instead of listening to the car as a car, I could focus on the specific quality of how the engine on a Ferrari sounds, for instance, or the auditory characteristics of the high-pitch engine whirring inside of a canyon. While this distinction has a history in the literature, it has run into issues as well.

In contrast to positing a difference between everyday listening and musical listening, an alternative would be to place the crux for different types of listening somewhere else. Following Nicola Dibben (2001), I suggest we replace the musical vs. everyday distinction with a distinction between listening to acoustic attributes of a sound vs. listening to sound source specifications, respectively, Both types of listening occur in musical and everyday contexts. For instance, earing the sound of a guitar, hearing an A chord, and hearing the start of a specific song are all different ways of describing the same auditory event.

Dibben (2001) provides additional empirical evidence in favor of shifting this distinction. In two related studies, Dibben asked participants to (1) categorize the similarity and difference between musical and everyday sounds and (2) provide commentaries on those sounds. In the first study, the question was whether or not subjects are more likely to group sounds together based on acoustic similarity or different kinds of similarity. Drawing on a pool of forty-eight sounds, participants were given 24 groups of three sounds during each group. For each triad of sounds, two conflicting pairings were presented: two of the three sounds were related by acoustic resemblance while two of the three sounds were related by common source features (e.g. Dibben (2001, p. 168) uses the example of a triad made up of Frying, Heavy Rain, and a Stream). Both

experts and novices showed a high-propensity to select matches based on acoustic features rather than source specification. They were nevertheless attuned to both the source of the sounds and the what the sounds specify, especially for musical examples.

In the second study, Dibben asked subjects to describe the sound samples that they had compared in the first experiment. She found that, when details of the context were limited, most participants had a general "default" mode of deferring to acoustic characteristics as the basis for comparing and contrasting among different sounds. Listeners referred to a wide variety of characteristics to explain their choices, including physical source and genre as the most prevalent two choices, followed by acoustic characteristics, then several more limited cases of elements such as emotion, function, and social context. Drawing these two studies together in conclusion, Dibben explicitly considers an ecological stance and argues that "meaning arises from the mutuality of object and perceiver, and sounds specify meanings and values for particular listeners, some of which can be mobilised at particular moments" (2001, p. 184). Considering the overlap of everyday and musical forms of listening, we will now turn to bring back Risjord's account of improvisation and the three main components of attunement, affordances, and meta-cognition.

Attunement and Ecological Information

Since listeners can interpret the same sounds in a variety of ways, what exactly do we listen to when listening to a sound? While I shall consider different ways of understanding sounds as affordances below, we first find that, regardless of the specific ontology of affordances one accepts, it is commonly acknowledged by proponents of ecological psychology that we live in a well-structured environment (Chemero & Turvey, 2007). An important feature of well-structured environments is that they have *invariant* aspects. The movement of soundwaves through a

medium may be chaotic from a certain vantage point. But it is not completely random. Likewise, light not only travels at a consistent rate but it bounces off different substances in similar ways across different contexts. As a result, perception in any sense modality becomes primarily about learning what these invariant structures are and attuning to various aspects of them when needed, either by choice or by necessity.

Perceiving invariant structures furthermore fits nicely into this account, since invariant connection between organism and environment decreases the need to posit robust kinds of internal processing and, perhaps, removes the need for internal processing entirely (Michaels & Palatinus, 2014). Instead, and in contrast to the cognitivist assumptions, perception is based on picking up information already available in the environment. The organism, in other words, simply has to attune themselves to that information correctly to do so.

An important aspect of attunement concerns the nature of information to which an organism is attuned. In ecological psychology, this information is explicitly ecological in kind. Ecological information relates to an organism's niche, which can be understood as the particular part of the environment in which an organism lives. A niche can furthermore be taken as a combination of a medium and aspects of the environment (Bruineberg, Chemero, & Rietveld, 2018).

A classic example of this combination between organism, medium, and environmental aspects can be seen in optic flow (Gibson, 1979). For our current focus on sound and music, however, there is an important parallel to optic flow in the case of an acoustic field (Gibson, 1966, chap. 5). An acoustic field is comprised of two elements: *wave fronts*, which specify the location of a sound, and *wave trains*, which specify the source of the sound. Both wave fronts and wave trains are impacted by the medium in which a sound is moving along with the

receptors available to any agent that hears the sound.

For a concrete example of an acoustic field, consider clapping your hands to get someone's attention. Sound waves are produced from the clapping and propagate outwards into the environment. The hearer can use clues from how these wave fronts hit their ears to localize the source, in turn. A louder sound in the right ear instead of the left will imply that the sound source is on the right, and therefore the person clapping is somewhere to the right side. At the same time, waves are not just useful for the location of the sound. Every sound wave will have wave trains that carry information, such as the fact that this particular sound was created by clapping hands. The same sound waves and wave trains simply would not come from a ringing bell or a knock on the door. At the same time, while people may clap differently, there are certain invariant features of the sound of clapping that holds across different cases.

We are now faced with the issue of the best way to understand these invariances in perception. Early and still influential accounts in ecological psychology (e.g. Turvey, Shaw, Reed, & Mace, 1981) posited that invariant features of ecological information are lawlike relations. For instance, based on features of what makes parts of the environment climbable, we can make generalizations about what is required for climbablility as an affordance in an organism's niche. Even after granting that many aspects of how one climbs may be culturally specific, plus the fact that there are times when additional states of the organism, such as fatigue, will have an impact on whether or not something appears climbable, the core point of something lawlike about the relation between parts of the environment and parts of the organism in the climbing process remains. It stays so much so that studies have indicated we can calculate the relative distance between a height in the environment and leg-length specifies to find whenever something, like stairs, affords climbing and when they do not. This so-called *tau* variable is a

ratio that holds across a multitude of different leg lengths and sizes (Warren 1984), and it tracks the invariant relationship between leg-length and height of the object to be climbed.

There are conflicting theories of what makes an explanation lawlike (Carroll, 2016) Generalization, in itself, is not enough for the whole story. A common addition to is that lawlike explanations operate in counterfactual terms (Lange, 1993). According to proponents of counterfactual accounts, law statements and lawlike explanations can be made with respect to a plethora of possible states of affairs, in contrast to accidental generalizations which cannot.

An additional feature of lawlike explanation in ecological psychology concerns a one-to-one correspondence between an aspect of the environment and the organism that perceives it. Such correspondences are primarily dispositional in nature, and Michael Turvey further suggests that these laws are "invariant relation[s] among substantial properties of things" (1992, p. 177) When the correct combination of substantial, dispositional properties is taken up in the correct context, the law will be actualized.

With these points in place, we can now reconsider how attunement relates to Hawkins' performance. Taken at a purely acoustic level, the attunement that Hawkins goes through is no more or less than what happens when I attune to the sound of hands clapping nearby or cars driving in the distance. However, in contrast to simply hearing claps and cars, the various sounds in Hawkins musical environment, from the bang of the drums to the comping of the piano and bass, cannot be understood without reference to his' particular skills with the tenor saxophone. A person who is listening to bebop for the first time, an experienced listener, and an expert performer all hear the same pool sounds when listening to *Body and Soul*, yet they do not attune to the same parts of those sounds.

There is here, however, a potential concern with lawlike accounts in ambiguous or

uncertain environments. Most of Gibson's examples assumed that an organism was operating in a standard environment in (at least semi-)optimal conditions – a sunny day in the case of vision or a quiet room in the case of acoustics, for example. Taken from a certain perspective, however, it's possible to say that the environment in jazz performance is uncertain. Moreover, not only does the environment underdetermine the sorts of choices Hawkins could make across different performances of *Body and Soul*, but it leaves open a space of uncertainty within any specific performance. There is nothing intrinsic to a particular chord that specifies an invariant, lawlike connection between what is heard and what is to be performed next, after all.

These considerations push us to account for ways that ambiguous and uncertain stimuli can be understood here, especially in relation to lawlike explanations of the counterfactual, 1:1 kind highlighted above. For an additional example of ambiguous cases, consider boarding and sitting on a train while looking out the window and daydreaming. The train slowly appears to start moving until you realize that your train has remained stationary the entire time. Such ambiguities fortunately don't contradict the claim that the environment is normally well-structured, however they raise additional concerns that I shall turn to address.

An appeal to mutualism (Clarke, 2005) may be a main option to salvage the lawlike account of ecological information. According to proponents of mutualism, the various sorts of ecological information simultaneously being tracked by an organism are not mutually exclusive. As noted above, for instance, people can hear the same sound source in different modes of attending. Based on considerations of mutualism, it is possible for multiple interpretations of the same source material to occur without being contradictory. These specific choices will vary depending on the skills and background knowledge of the perceiver, as well as the nature of the current task at hand. In addition, there is a further point to be made about a distinct difference

between unclear information, such as the case of listening to a song in a noisy environment, and uncertain information, such as what a jazz player will play next in response to the song as already played. What Hawkins is faced with is clear but uncertain information.

While the previous considerations were largely in the realm of perception, the use of perception in these cases is tied into the sorts of decisions and actions made during the performance. The problem with either of these previous moves, especially paired with the idea that no internal processing is needed in perception, is that we are not ultimately given an account of why or how Hawkins decided to play one particular run instead of another during his solo. Furthermore, there may be a 1:1 correspondence in the realm of acoustics that makes sets of sounds the same across different contexts, but there is no such lawlike correspondence that will specify any connections among the notes played and future notes to be played. The decisions made during a jazz solo, in other words, are many-to-one, not one-to-one.

So much for the strict lawlike view of information. What about other alternatives? In response to these concerns, I turn to follow insights from Jelle Bruineberg, Anthony Chemero, and Erik Rietveld (2018) on the nature of ecological information. According to Bruineberg, Chemero, and Rietveld, lawlike ecological information tracks "a lawful regularity in the ecological niche between structure at a point p in a medium and an aspect of the environment at point q such that there is a 1:1 specifying relationship between p and q," while general ecological information includes "any regularity in the ecological niche between different aspects of the environment (x and y) such that the occurrence of x makes y likely" (2018, Sect. 2). Ambient light carries lawlike ecological information. A clock, in contrast, carries general ecological information. There is no lawlike, 1:1 relationship between the time and the hands on the clock (e.g. a clock could be broken, fast, or slow), yet an operant clock nevertheless caries some kind

of ecological information by allowing us to create and respond to various regularities in the environment. Likewise, on this account, the two types of information are not opposed. Lawlike information is instead a subspecies of general information.

Jazz performance, as with other types of music performance, will draw on elements of both lawlike and general ecological information. An experienced trumpet player performing *Straight, No Chaser* will know that blowing their horn and pressing keys in a certain configuration will bring about specific notes. They also know that adding a mute will change the quality of the produced sounds in specific ways. Both of these facts track lawlike ecological information. At the same time, the chord changes of the song that constrain any and all musical decisions made during the song do so without the same sort of strict correspondence as the first two examples. I can predict ahead of time how a trumpet will sound when played with a straight mute; I can only predict that a solo played within these chord changes will likely include these notes and not these others.

As a result, what general ecological information adds over and above lawlike information for understanding jazz performance is that it gives a ground for ecological information that stands between a strict law and a chaotic situation requiring robust internal processing by the agent to make sense of the world. It addresses the specification worry by allowing us to trace the regularities of the music itself, along with attendant social and performance norms, that make certain choices constrained and more likely, while they still remain open to differences across time and contexts. We can also apply these ideas with an aim to interpret a piece of music from an external perspective and/or to perform the same piece of music as a musician.

To conclude this section, I will consider an additional problem of tracing lawlike relations in the case of instruments that maintain some distance between what produces a note

and the final output, usually to be found in the case of electronic or MIDI instruments. The difference here would be between, e.g., playing an acoustic drum and playing an electric drum. While I can coax out a variety of distinct sounds from an acoustic drum, I am limited by a variety of constraints for the sorts of sounds I can make. The electronic drum is not unlimited in scope, but I can make it moo, quack, ring, ding, or sound like a well-tuned snare drum in ways that simply cannot be done with an acoustic drum.

While one could even take the stronger claim that the rise of technology has divorced sounds from an ecological context as well as their ecological origins, that would be incorrect for at least two reasons. First, in the case of MIDI instruments, their purpose is often to emulate sounds from other instruments. Doing so will requires properly tracking various (invariant) timbre and sonic structures that make a flute different from a tuba, for instance. As such, these sorts of devices utilize invariants like the painter can use invariants in their 2D paintings to express depth and contrast on a flat plane.

Second, even in cases where there are no robust, 1:1 invariant present, we still hear a sound as noise. Noise, moreover, is a complex phenomenon that can be further subdivided into different kinds, including white noise, pink noise, blue noise, and violet noise, among others. In other terms, the drone of a chant is not the same as the static from a TV set. As a result, while noise may not have the same depth of invariant sorts of structures as other sounds do, that is not to say that they fail to either be meaningful sorts of information within the environment. Noise may also have a direct impact on the transfer of acoustic information in ways that have been measured in ecological acoustics (Namba & Kuwano, 2004).

In this section, it has been suggested that a musician attunes to both lawlike and general ecological information. What it does not yet provide is a full account of how that information is

connected to choices and actions made during performance. To answer that issue, we will now return to affordances in more detail.

The Ontology of Musical Affordances in Bebop: Dispositions or Relations?

Following Edward Reed (1996, p. 40-41), asking 'what are affordances?' can be taken at either general/abstract or specific/concrete levels. Specific affordances require us to look at the concrete implementation of a particular affordance to a specific type of organism. Take the case of a human engaging with the affordance climbable, for instance. Studying this specific affordance requires considering how stairs afford climbing to a bipedal human. In a similar manner, an example of a specific musical affordance, following the work of Joel Krueger (2014), is that music affords entrainment for human animals. The sorts of entrainment he considers scale from an infant bouncing along with a beat to a trained musician that can leverage entrainment, musical instruments, and their skills to sustain complex performances. Furthermore, the story about the specific affordances of climbing and entrainment would be different if we were considering how an elephant or snake climbs, for the former, or how a cockatoo entrains with a beat, for the latter.

General cases of affordances abstract away from some particulars to consider how an affordance holds across morphological, biological, and social differences. Bipedal climbing is different from quadrupedal climbing in concrete ways, after all, but we can nevertheless speak of the climbable affordance across such differences. In regards to musical entrainment, we can consider the general ability for entrainment to music across species.

In both general and specific cases, being an affordance necessarily entails a relationality between organism and environment. We considered some aspects of this relationality in the previous sections. In regards to how an organism attunes to their environment, relationality is

essential since we find a well-structured environment, on one hand of the relation, and an organism that is able to track and take advantage of that environment, on the other hand.

How to understand this relation, both in regards to the organism and environment, is the core ontological issue of affordances that I will consider in this section. There are two main alternatives for the ontology of affordances. The first alternative is dispositional accounts. Proponents of these accounts suggest that affordances are *dispositional properties* of objects in the environment that, when paired with *effectivities* of an organism, result in action. The second alternative is relational accounts. Proponents of these accounts suggest that affordances are *higher-order relations* between *features* of the environment and *abilities* of an organism. In what follows, I shall argue that relational accounts do a better job of describing musical affordances in jazz improvisation. I shall also utilize musical entrainment as my main example of an affordance, since it is an essential part of musical performance.

Dispositional Accounts

One of the first formal definitions of affordances was a dispositional account proposed by Michael Turvey (1992). This account had three primary features (1992, p. 178): First, a disposition exists prior to its realization. Just like salt has a disposition to dissolve without being placed in a liquid, the steady pulse of music that affords entrainment exists independently and prior to any entrainment. Second, insofar as dispositional and causal propensities come together in a pair, dispositions entail an intrinsic kind of complementarity between organism and environment. In order to act on an affordance, an organism needs effectivities that pair with environmental dispositions. In the case of entrainment, relevant organism effectivities include ears to hear, bodies to feel, and mechanisms that help the musician or dancer maintain the

entrainment once it has begun, among other things. Third, the combination of a disposition and its actualizing circumstance will always result in the actualization of that disposition.

At first pass, the third criterion seems most difficult to account for in the case of musical entrainment. We shall return to this worry again below as one of the major issues with dispositional accounts. However, and even though I ultimately suggest that this issue is a serious problem, there are two *prima facie* points in favor of a dispositional account here, at least in regards to musical entrainment. First, it's possible to entrain to music without being consciously aware of being entrained. This result is important since there may be cases where a person does not report feeling entrained to music when they are, at least to a certain extent. Second, while it is possible to consciously counteract being entrained to music, this is only possible after entrainment has occurred. Moreover, the final criterion of necessary actualization only requires that a disposition will occur in the right contexts, it does not have anything to say about any length of time it must be maintained after it has begun.

We can provide further clarification for Turvey's definition by additional comparison to dispositional properties in non-affordance cases. For instance, let's consider the common dispositional property solubility vis-à-vis salt. In light of the three criteria, we find that (1) Salt must be soluble before it is actually dissolved, and it must have that property in order for us to talk about solubility in the first place. (2) While this dispositional property existed prior to any concrete realization of salt dissolving, it only does so because there are substances that are able to dissolve the salt.[5] (3) If the conditions are met, salt will always dissolve. If it were basically chance that salt dissolves when placed in a glass of fresh water, for instance, we would have

[5] This formulation may beg the question in an important debate about whether or not dispositional properties can exist without instances of those properties in the world. For those who reject a Platonic approach to dispositional properties, an additional caveat can be added that dispositional properties presuppose realization in the real world before we consider them *bona fide* properties (for more, see Heras-Escribano, 2017).

reason to reject solubility as a property of salt.

Two of the strongest criticisms against the dispositional view can be found in the work of Anthony Chemero (2009) and Silvano Zipoli Caiani (2014). Chemero's criticism (see also Stoffregen, 2003) focuses on the third point of Turvey's definition concerning the claim that a combination of effectivity and affordance will automatically result in action. Caiani's criticism focuses on how to mesh the idea that the dispositions underlying affordances are properties that are necessarily geared towards concrete action and the possibility of affordances being seen in non-actionable items, such as paintings or pictures.

According to Turvey, "Dispositionals *never fail* to be actualized when conjoined with suitable circumstances. Disposition and suitable circumstances equal actuality" (1992, p. 178, emphasis added). The problem with such a strong statement about always actualizing an affordance is two-fold. First, any situation will be rife with various affordances for an organism. A piano can be played by pressing down the piano keys, but it can also be sat on, rolled to a new spot, or chopped up, to name just several possible affordances. If we take Turvey at his word, however, once a person stands face-to-face with a piano, all of these would *necessarily* follow. Such a result is absurd from a practical position and poses both empirical and theoretical problems. As Stoffregen has noted, "in any given situation many actions are possible, but the great majority of them do not come to pass; that is, they are not actualized…Affordances are what one can do, not one must do" (2003, p. 119). There are likewise times where the ability to ignore an affordance may be as important as tracking and acting on available affordances.

Turvey may reply that this criticism understates the role of circumstances in establishing what counts as properly conjoining a disposition and an effectivity. I grant that this reply gets around some of the worst concerns highlighted in the previous paragraph. Nevertheless, it results

in the second worry that being in the correct situation will necessarily bring about an action. While it may be true in some cases, such as the case of unconscious entrainment highlighted above, it doesn't hold in others. This breakdown is particularly noticeable for musical improvisation when musicians intentionally refuse to take advantage of affordances across different yet similar cases. Moreover, if a player did realize the same action pattern over and over, they would fail to live up to essential features of bebop performance practice.

Moving to Caiani's concerns, a core issue becomes that dispositional accounts are flawed since there are cases when affordance perception can occur without the existence of "action-related dispositional properties" in the environment (2014, p. 286). This issue, in turn, is that affordances are perceived in cases both where a physical object is present (e.g. when the physical piano is sitting in front of me) and in cases where it is only represented by another means, such as a painting or picture of a piano.

The fact that we cannot perceptually distinguish between the properties of a real cup and those in a photo is a problem for dispositional accounts. The problem, however, is not that a person cannot tell important differences between a picture not affording drinking in the way a physical cup actually affords drinking. Most people will not make this kind of mistake in their everyday lives. The problem is instead that, if affordances are equally perceivable in both the physical object and mere representations of that object, then it is not clear that the affordances are dispositions of the physical environment.

The defender of a dispositional view will not be saved here by the move to circumstances. The reason they cannot do so, as noted by Caiani, comes about in relation to the mix of what he calls the "actualization requirement" (i.e. the third criterion from Turvey's definition canvased above) and the "identification requirement." This latter requirement is that

"Perceiving an affordance involves the presence of a property bearer suitable for action" (Caiani, 2014, p. 280). While not identical to Turvey's second criterion as I have discussed it so far, we can grant that the realism of his account, paired with the discussion of how dispositional properties work in general, provides enough support for granting the identification requirement as part of a dispositional account of affordances.

With these two requirements in place, Caiani runs his argument by appealing to the empirical fact that affordance perception occurs even if there is no object in the environment as the bearer of those dispositional properties (2014, p. 288-9). In his words, "as the property of being soluble does not apply to an abstract object, say a mathematical law, the property of being graspable does not apply to a two-dimensional object" (2014, p. 288). Yet since there is evidence that at least on a neural level the brain responds in similar ways to both pictures of cups and to actual cups, the dispositionalist account seems to be in trouble.

Before turning to relational accounts, however, I will first consider a distinct, non-factualist dispositional alternative that has been developed by Manuel Heras-Escribano (2019). The main difference about this account and Turvey's version is that it takes a notion of dispositions based primarily on Gilbert Ryle's understanding of the concept. In contrast to single-track dispositions such as solubility or brittleness, Ryle (1949, p. 32) points out that mental dispositions can be realized in a many-to-one format. He furthermore argues in favor of non-factualist dispositions, which includes skills, tendencies, and habits of an organism. Both of these components offer a potential response to the previous worries.

In regards to Chemero's concern, the many-to-one nature of correspondence for Rylean dispositions allows that there would be many ways to act, or not act, on a given affordance. Thus, it is possible that an affordance will always be realized when complemented with an

organism's effectivities, but those realizations will look different even in similar situations and for the same organism (e.g. a piano may be played in many different ways in a concert setting). In relation to Caiani's worry about the actualization requirement, since the relevant dispositions are here non-factualist, they allow for an appeal to the habits of an organism as an essential feature to understand affordances. In turn, a possible reply to Caiani's main argument would be that it is, in fact, no surprise for affordances to be perceived in pictures as well as for physical objects since, in the former case, it's simply the activation of a habit that tracks a similar invariant structure of a cup, even if that cup is not there to actually be grasped. We have habits and practices of holding and grasping mugs, after all, and certain activations of those habits can happen in either case.

There is, however, a distinct downside that goes along with the positives discussed for Heras-Escribiano's version of affordances. The main concern is a dilemma for the dispositionalist that may collapse their position into a relational approach. Taken from the observer's perspective, non-factualist dispositions operate as an "inference-ticket" that allows us to make claims about how an organism is currently behaving or would behave in certain contexts. As a result, in Heras-Escribiano's words, "Dispositions are not discrete entities to which we point with words, but terms that allow us to infer more detailed statements and to offer richer explanations of the behavior of organisms and objects" (2019, p. 84). This shift from metaphysics to semantics makes sense for Ryle's overall project of rejecting Cartesian dualism and its attendant metaphysics, although it's not clear that it is the best option for a defender of the dispositional account to take.

To see why in more detail, we first find that these considerations lead Heras-Escribiano (2019) to conclude that making the claim 'something affords an action' is to say:

> That I have the tendency to do this or that action when I encounter certain aspects of the environment under the right circumstances. Affordances are not, then, extra entities that proliferate in our world, but a way to explain the complementarity of organism and environment, including the readiness of certain agents when they encounter certain objects in our environment, such as when I encounter objects that I can grasp or throw. This throwability or graspability is not an extra entity beyond an agent and its surrounding environment, but a way of making sense of the complementarity of the shape of the cup and the features of my hand. (p. 87).

While this claim clearly suggests that affordances are not entities in their own right, it does not yet establish where affordances are located. There are three main options for location. The first is to place them in the environment, as Turvey does, since Gibson was clear that affordances are what the environment affords an organism. With a focus on abilities, skills, and habits, however, it doesn't seem possible that non-factualist dispositionalism could go this route, unless Heras-Escribano was willing to grant that the environment has skills and habits just like an organism.

The second option, which follows from the idea that skills, abilities, and habits are better understood as features of an organism, would place affordances in the organism. After all, Ryle's dispositionalism is about and within agents. Unfortunately, affordances are not going to be primarily located within the organism, since doing so would get the relationship between organism and affordances backwards. The third and final option is to place affordances in the agent-environment system. While this is the best route, both in itself and by process of elimination, it collapses non-factualist dispositionalism into a relational account.

In addition to these two types of dispositional accounts and their flaws, however, there is one final argument I will consider in favor of dispositionalism. According to Naoya Hirose

(2011), a major source of conflict between proponents of relational and dispositional accounts comes from different standpoints that one could adopt towards affordances. The first option is from the point of view of the observer or experimenter and the second is from the point of view of the organism. In addition, Hirose argues that we should focus on cases of what affordances are from the organism perspective, both since it is more related to Gibson's original reasons for postulating affordances in the first place (Hirose, 2011, p. 235) and understanding them from this perspective will provide a better theory of action.

Hirose then argues from the organism perspective that only properties of the environment matter. Capabilities and skills, in contrast, are fairly irrelevant to an organism perceiving an affordance, since such skills are "relatively constant" and act as "the frame of reference" that is taken as given through the perceptual process. In turn, she argues that the relational account fails to give weight to the inherent asymmetry between an environment and organism, wherein the environment *affords* something *to* the animal.

One response to Hirose would be to claim that her account conflates an epistemic difference into an ontological one. Both dispositionalists and relationalists accept that there are important differences between the phenomenology of affordance perception for an organism, on one hand, and an "objective" account of what affordances are from a third-person perspective, on the other. This distinction furthermore can be run parallel to Reed's distinction between general and specific affordances. In either case, whether or not abilities are "the frame of reference" for an organism does not have any relevance for what an affordance is, even if it is highly relevant for understanding how we come to know about affordances in the world.

Relational Accounts

In contrast to dispositional accounts, the main alternative takes affordances as relational features between an organism and its environment (Chemero, 2003; 2009; Stoffregen, 2003; Rietveld & Kiverstein, 2014; Caiani, 2014). Stoffregen (2003), for instance, offers a more formal definition for affordances as a higher-order relation of the animal-environment system.[6] Such a relation does not entail we understand affordances as a new entity of the environment. Furthermore, while Turvey's dispositional definition places affordances in the environment (albeit only realized in action when paired to an organism's effectivities), the relational definition places affordances at the level of the animal-environment system.

While more pros and cons could be said about Stoffregen's account, I propose that we use Chemero's (2009) definition of "affordances 2.0" as our primary touchstone of relational accounts. Affordances 2.0 takes affordances as *relations* between the *abilities* of an organism and *features* of the environment. According to Chemero, Affordances 2.0 emphasizes the dynamic coupling between an organism's abilities and its environmental niche (see also Heft, 2001) that, in turn, carries several benefits over the dispositional accounts canvased above. These benefits include moving the definition of affordances closer to working practices of ecological psychologists and bringing together insights from ecological psychology with other branches of 4E cognitive sciences.

While more can be said about these two benefits, a third upshot is even more important for our current purposes. In Chemero's words, "this reconceptualization of affordances is a variety of niche construction that occurs over shorter time scales and in which the constructed niche is an animal's individual behavioral, cognitive, and phenomenological niche" (2009, p.

[6] Stoffregen furthermore distinguishes between affordances and behavior, such that the former is necessary for the latter, but behavior only comes about in conjunction with some kinds of goals or intentions.

151-152). Niche construction (Laland, Matthews, & Feldman, 2016) is an idea from evolutionary biology, according to which organisms are influenced by their environment at the same time as they actively build and maintain their niches. As a result, a kind of circular causality is developed in both the case of niches and affordances (Ryan & Schiavio, 2019).

In these three points about affordances 2.0, we are able to start seeing how a relational account is useful to understand jazz performance. Jazz itself can be understood as a type of niche, and the affordances in that niche are as varied as the wide array of affordances in any environment. Considering the dynamic nature of niche construction, moreover, the relational account offers the conceptual tools to better understand how affordances change over time and help better develop an account of sociocultural and musical affordances.

There are several counterarguments to relational accounts. One of the main ones concerns the use of relations as the basis for affordances. A common example of a relation would be different heights, such as taller-than or shorter-than. In these cases, we can only make sense of the relation if we have (at least) two points of reference. While I could be tall, I cannot be taller-than unless there is another person or thing that I am, in fact, taller-than. From these, Heras-Escribiano argues that relational approaches fail to account for affordances concerning "actualization, potentiality, or change in the current situation" (2019, p. 87).The sense of actualization is lost for relational accounts, according to Heras-Escribano, since they don't necessarily entail changes in the environment, and change in the environment is required for action. Likewise, since affordances are fundamentally geared toward action in the world, the fact that they could exist independently of change appears to be a flaw with relational approaches.

In response, I suggest that the heart of this concern mirrors the mistake of rejecting a dispositional account while only focusing on simple, one-to-one dispositions. Being taller-than is

clearly not an affordance, even though it is a relation. Affordances *qua* relations are relations geared towards action; the cup affords picking up because it is the right material, shape, and weight, plus the fact that it contains something I aim to drink or there is a target at which I aim to throw it. It is these sorts of relations that make actions possible, not just any relations that happen to be in the world. Finding examples of relations that don't serve this purpose are therefore moot as counterexamples to the relational theory of affordances.

A second concern for relational accounts is how to conceptualize expert skills and the existence of particular affordances. Following Eric Rietveld and colleagues (Rietveld & Kiverstein, 2014; Rietveld *et al.,* 2018; van Djik & Rietveld, 2018), who developed an account of affordances within a larger Skilled Intentionality Framework (SIF), we can address this concern as well as further flesh out the dynamic nature of affordances by establishing how a subset of affordances becomes salient among the set of all available affordances. Rietveld and colleagues suggest that the distinction between a *landscape* of affordances and a *field* of affordances will help explain this process of foregrounding. In addition, Rietveld and Kiverstein (2014) introduce the Wittgensteinian idea of a "form of life." The form of life can be taken across various levels, including the form of life of a species (e.g. human, dog, cat, moose…), the form of life of particular cultures (e.g. Thai, Egyptian, Greek, Irish…), and the form of life of particular social groups (e.g. musician, teacher, politician, electrician…).

The landscape of affordances includes all possible affordances available to an organism, while the field of affordances captures those that are salient to them at a given moment. Moreover, affordances in this sense are not tied to individual organisms, which marks a slight shift from Affordances 2.0. Rather, affordances are defined as the relation between aspects of the environment and skills found within a form of life. This shift helps explain how general

affordances – e.g. stairs are climbable - can exist without an individual being present in a given environment, and how specific affordances – e.g. a book is readable - develop in relation to the concrete practices of both individuals and groups.

Let's consider the case of musical instruments as an additional example. An electric guitar affords plucking, strumming, tuning, bowing, stringing, burning, and swinging, among many other action-oriented opportunities. All of these affordances will be present in the landscape of affordances for an organism who is capable of engaging with them. Jimi Hendrix and I could, at one time, both strum a guitar, but there are many strumming affordances available to Hendrix as an expert that are unavailable to me as a novice. As a result, our respective fields of affordances vary. In addition, the affordances that Hendrix had originally learned and popularized are now available within the form of life of guitarists, even though he has long since passed away.

Another aspect of affordances in the SIF concerns why a particular affordance appears to solicit action or other action-related responses. The key factor here is the salience of the affordance. Within the landscape, a field is constructed through variable saliences of different affordances. These affordances will be more or less salient for a variety of reasons. Sometimes features of the environment increase or decrease the salience of an affordance, such as when a novel chord phrasing draws the attention of co-performers or an infant entrains to a strong musical beat. Other times social factors impact it, such as the case of an approving or disapproving look from another musician in the group or social expectations about what an audience likes to hear in a song. Still other times salience can be negotiated by how an individual consciously centers their attention, such as whether or not a soloist pays more attention to the main melody or the comping of the rhythm section. Whatever affordances are actually salient at

a moment will be part of the field of affordances. The landscape of affordances for Hendrix's guitar always afforded burning, for instance, but it only entered his field of affordances during the performance in Monterey in 1967.

We can also apply the field/landscape distinction to Hawkins' recording of *Body and Soul*. As noted above, Hawkins' took this standard, as many bebop players take a standard, as a series of chord changes over which to improvise his solo. This choice was most notable in the approach of the comping musicians, who provided a sparse musical landscape behind him that ultimately highlighted these changes without much reference to other features of the song (especially in contrast to the Armstrong version discussed above). In making these musical choices, the band helped to stabilize a particular field of affordances within the broader landscape offered by the song.

Furthermore, in relation to the form of life of tenor sax players, this particular performance marked a new set of possibilities beginning to unfold within it. These changes cannot be distilled to the genius of Hawkins' alone, of course, although the fact that it was him making these decisions at the time is not an irrelevant consideration to keep in mind. As a result, he not only impacted the particular landscape for his own performance, but for the overall landscape of future performances as well.

In this section, we considered two alternative ontological approaches to affordances: dispositional accounts and relational accounts. Dispositional accounts either ran into serious internal flaws or collapsed into relational accounts. I then discussed a series of additional considerations about how affordances become available to individuals during action. Thus far, however, besides one aspect of availability being acknowledged as self-produced, I have said little about what happens inside of an organism when they attune to affordances. Part of this

decision was to respect the assumption that there is little to no need for appealing to internal processing when it comes to the dictates of ecological psychology. However, this position does not seem tenable in a strong form moving forward. Instead of rejecting any and all internal happenings as relevant for an ecological account, there must be some role for the brain and body to be playing beyond merely being dragged along from one solicitation to the next. What this role is exactly will be the main topic in the following section.

Meta-Cognition: The Brain as Resonance Organ

Orthodox ecological accounts reject any appeal to internal processing as part of the cognitive story, especially in any guise that supports computational or informational processing approaches to the mind. As noted explicitly by Claire Michaels and Zsolt Palatinus as their 9th commandment of ecological psychology: "thou shalt not make unto thee any mental image or likeness *of anything*" (2010, p. 25, emphasis added). Under this assumption, instead of looking to the inside of an agent, understanding the link between agent and external world is all we need to tell a full cognitive story; the well-situated agent will pick up invariant information from the environment, either by choice (e.g. in the case of self-driven actions) or not (e.g. when something from the environment solicits and grabs our attention in the moment), which will in turn drive iterant loops of perception and action. Likewise, in neither the case of an agent using the environment or an agent being pulled along by the environment will they need to add something extra to external information picked up in the perception-action-cognition process.

At the same time, even the devout defender of ecological psychology recognizes that plenty of stuff is happening inside an agent during cognition. Most notable among these happenings is often what is going on in the brain, although the full, embodied story will require taking into account other bodily aspects in cognition, such as activity in the gut guiding behavior

both in concert and independently from brain activity (Davidson, Cooke, Johnson, & Quinn, 2018; Liang, Wu, & Jin, 2018), among others. As a result, internal processes need to be better understood in relation to ecological aspects, since no account of cognition would be complete without both sides.

This second point, in turn, prompts the question of how we should understand what internal processes are doing vis-à-vis the usual explanations provided by ecological psychologists. Or, more narrowly, we can ask what exactly is the brain doing within an ecological account of cognition? For Risjord, the answer to this question is that the brain is taking part in meta-cognition. As a result, what exactly meta-cognition looks like will be the focus of this section.

First, one way that ecological psychology has been applied in the context of music psychology and internal functions of an agent can be found in the work of Charles Nussbaum (2007) on musical understanding in Western tonal art music. Considering the clearly stated focus on Western art music and its listeners (2007, p. 38-40), this account cannot be directly applied to our current focus on bebop improvisation without some modifications. Nevertheless, most of the basic tenants underlying the theory can be applied without much reworking, such as a particular role being placed on "acceptable" moves during the development of solos that respect certain tonal, chord, and key related constraints. I shall therefore apply Nussbaum's account without much alteration below, flagging any immediately relevant changes if they are needed.

Nussbaum's account fuses together Fred Lerdahl and Ray Jackendoff's Generative Theory of Tonal Music (GTTM) with embodied cognition. This link is most clear in his claims that musical understanding is heavily grounded in the evolution of the human auditory system. When traced back phylogenetically, there is a strong overlap between the haptic system and the

auditory system. The cochlea, a central organ for audition, has a homological relation to the sacculus of fish, which is a tactile sensory system that also serve a rudimentary hearing purpose, for instance (Nussbaum, 2007, p. 52). There are also important functional relations (for example, auditory and haptic experiences unfold in a serial way over time) and structural relations (for instance, certain uses of both audition and haptic sensations are underpinned by mechanoreceptors, in contrast to other types of receptors prevalent in different sense modalities and different uses of auditory or haptic pathways) between the haptic and auditory systems.

In addition to the tight link between touch and hearing, there is another tight link between perception and action on this account, especially since listening to music invokes a variety of ways one can act with music. Following these two links and Gibson's definition of affordances, Nussbaum suggests that music itself can act as an external representation, a la musical affordances. It does so by being a series of invariant relationships that are intertwined with motor and action responses to the music.

If this kind of representation were the only kind in the account, it would be easily amenable to orthodox ecological accounts. External representations, however, are not the only sorts of representations for Nussbaum. He instead argues that musical surfaces are "a carrier or vehicle from which information can be extracted by performing appropriate transformational operations that are *supported by representations in the human mind-brain*" (2008, p. 23, emphasis added). These internal representations are the rules and general posits of Lerdahl and Jackendoff's GTTM implemented in the motor system.

While more could be said about how this implementation process operates, the details are not needed to motivate the main idea that, here, internal representations are taken by Nussbaum as a necessary part of explaining the role of the brain. Fortunately for ecological psychologists,

breaking one of the ten commandments is not the only available option to explain how internal neural processing links with affordances. In contrast to representationalism, a less heretical alternative is to appeal to *resonance*. Doing so may help address some of Nussbaum's concerns without appealing to internal representations as an essential part of the cognitive *explanans*.

The idea of resonance first made several minor appearances in Gibson's work, wherein he briefly suggested that we understand the brain as "resonating" with the environment. In his words (1966):

> Instead of supposing that the brain constructs or computes the objective information from a kaleidoscopic inflow of sensations, we may suppose the orienting of the organs of perception is governed by the brain so that the whole system of input and output resonates to the external information. (p. 5).

Later in the same text, Gibson clarifies a possible misreading of resonance by noting that "The 'resonating' or 'tuning' of a system suggests the analogy of a radio receiver. This model is inadequate because there would have to be a little man to twiddle the knobs. A perceiver is a *self-tuning* system" (1966, p. 271). For our immediate purposes, what's important with this claim is that the resonance will be shaped by aspects of both the environment and the organism. Furthermore, embracing the idea of the brain as a resonant organ does not presuppose that the brain is only controlling a passive organism since resonance is not always itself a passive process of information pick-up.

While the concept of resonance in ecological psychology has been around for over 50 years, it remains rather underdeveloped. Furthermore, if resonance is taken as the claim that action is driven only by an organism resonating with the environment, we run into a problem with the case of jazz improvisation. This problem, as discussed in relation to attunement and affordances above, concerns the fact that even informationally rich invariant features of a performance still leave a wide latitude of possible choices for musicians to make while playing a

song. In other words, a musician may resonate with parts of the sonic world during a solo, yet their choices are not overly constrained by the resonance process. This opening between the two leaves a space where one may assume that some kind of internal representations must play a role.

One possibility to diffuse this particular concern is to place the self-tuning brain back at the core of the picture. Yet highlighting self-tuning in resonance alone does not solve an additional, related issue about the nature of agency. Put succinctly, if the central claim of resonance is that all the brain is doing is resonating with ecological information, then no additional processing should occur inside the organism. But some of the choices made by the musician extend beyond information available in the environment. There is thus, it seems, a need for internal processing that cannot be explained solely by appeal to resonance. Thomas Fuchs (2018) has offered one account that considers the idea the brain is a resonant organ, and to which I now turn to see if it can be used to address this worry.

Drawing directly on a metaphor of jazz improvisation, Fuchs claims that "the brain is not the conductor of the body; rather, it is like a musician in a group of jazz musicians jointly improvising on the basis of certain chords" (2018, p. 134). Fuchs likewise highlights that the notion of resonance comes from considerations about acoustics and oscillation, making it ripe for application in the case of music performance (see also Shepard, 1984). In contrast to traditional representational accounts, which take a static view of representations in relation to action and perception, Fuchs suggests that this acoustical focus brings the temporal nature of cognition to the fore. As a result, in his words, "Resonance contains a dynamical as well as a rhythmical element and thus establishes a *temporally* overarching relation between the systems involved…'resonandum' and 'resonans' thus cannot be separated" (2018, p. 166). Such

inseparability is furthermore taken as a sign that the explanation of resonance, and of the brain in general, cannot be representational in nature.

Resonance can further be understood on two levels for Fuchs. First, the brain and body resonate with each other in a dynamical, intertwined process of homeostasis. This level of resonance is tied into the fact that the brain is "*the 'integral' of the overarching process of life* which encompasses the whole organism" (2018, p. 119, emphasis in original), from the level of densely interconnected brain activity across the brainstem and cortex, to the role of affect as essential to cognitive activity, and out further still to the densely intertwined efferent and afferent feedback between the brain and non-neural body.

Second, there is a resonance between an organism and the environment. This resonance occurs through a "dynamic set of isomorphic patterns" that develop between the brain, body, and world. An example of isomorphic patterns would be when a specific (or similar) brain pattern occurs whenever in the presence of a specific (or similar) environmental context. Moreover, neural evidence has shown that brain patterns change in response to learning new habits and skills, such as the increase in musical ability resulting in different neural activations compared to novices first learning how to play (Oechslin, Van De Ville, Lazeyras, Hauert, & James, 2013). Such results suggest that the dynamics of these isomorphic patterns are skill dependent, as we would expect on the case of a relational account of affordances.

These two levels can furthermore be parsed out in relation to Vicente Raja's (2018, p. 33) distinction among three different target scales for resonance: (1) the agent-CNS (Central Nervous System) interaction, (2) inner-CNS interaction, and (3) CNS-environment interaction. While Fuchs' account takes on resonance in all three of these target scales without always prying them apart, following Raja, I suggest that the agent-CNS interactions is the target explanatory level for

our current case. At this scale, the focus is how activity of the CNS resonates with the activity of an agent within their environment. The inner-CNS scale, which tracks the first level presented by Fuchs, is an important use of resonance, no doubt; the problem is that it fails to track the ecological level that is an essential part of the main story of ecological psychology that we are exploring here. Finally, the CNS-environment scale fails to account for important ways that an agent is able to modulate and alter their interactions with the environment.

An immediate problem with Fuchs' analysis is that it may be too focused on matching between patterns of the brain and patterns of the world. Part of this concern may be an artefact of his focus on perception in non-improvised settings, where there is a clearer connection between the invariant structures of the environment and how an organism tracks them. The worry may also be the result of the artificial limitations from an account that is still undergoing development and additions. One additional avenue available for Fuchs to respond here is that the neural patterns he is considering are primarily patterns relevant for action, not just abstract isomorphic patterns. After all, the patterns being resonated with are from a rich affordance landscape, laden as it is with a variety of meaning for the organism. Thus, resonance is not about locking down static features but, instead, grasping patterns in the world that ultimately allow organisms to act on the rich, complex, and dynamic environment.

Unfortunately, while not entirely inconsistent with Fuchs' theory, some parts of Fuchs' analysis go against this reading. For instance, after appealing to Herbert Dreyfus and Charles Taylor on an Aristotelian view of the mind, Fuchs (2018) claims:

> the brain could be conceived as a matrix, which like the mind is able to 'receive all forms,' that is to say, to take them over in its own structure as neural patterns or potentials. In the actual perception 'mind and object become one,'

> corresponding to an encompassing resonant system state in which *the same pattern* or form is activated in the brain as it is displayed by the object." (p. 167, emphasis added).

In Aristotelian terms, the form of an object can be distinguished from its matter, and it is the form that is taken into the mind during perception. However, this idea runs into an additional problem here concerning exactly what it means to be isomorphic. What is the exact isomorphism between a good wine and its taste? Is the structural similarity of the isomorphism to be taken at a first level of isomorphism (e.g. the brain resonates with the invariant features for the taste of this particular wine) or a higher-order (e.g. the brain resonates with the invariant features of the experience of tasting this wine)? If taken in the first level sense, what does it mean to resonate with the "taste" of the wine? Indeed, which specific properties are actually being resonated with in cases of perception, aesthetic or otherwise?[7]

In addition to the problem of isomorphism, another way of capturing my overall concern with Fuchs' account is to return to the problem of environmental uncertainty during jazz performance.[8] Accepting uncertainty here is not rejecting the idea that jazz performance takes place in a well-structured performance environment. Instead, it highlights the fact that, in addition to the importance of an agent self-tuning their resonances, the music does not offer an overly strong boundary on choices made during performance. A soloist is not just a coupled oscillator resonating with patterns of their environment. They also must create improvised

[7] Nussbaum considers these two points in relation to a representational account in (2008, p. 28).

[8] This sort of conundrum should be kept separate from concerns about poverty of the stimulus. Ambiguity, in the sense I am using it, is also not addressed in Gibson's (1966) Chapter 14 on The Causes of Deficient perception. In that chapter, Gibson considered both cases where information is inadequate or psychological processing are insufficient. In contrast, the problem for its application in jazz is not a deficiency of perception but rather that the brain may be doing more than resonating or changing how to resonate in order to achieve the overall task at hand.

choices in the moment of performance. To explain the fluid, creative possibilities of jazz performance we must go beyond matching resonance to something more.

We have two options available for us on this front. The first is to return again to a representationalist alternative presented by Nussbaum, who drew heavily on this particular point from the work of Roger Shepard. Speaking directly in regards to ambiguity, Nussbaum notes that "degraded and ambiguous inputs immediately reveal the extent to which information extraction depends on pattern completion, which in turn requires internal representations and constructive procedures that operate over these representations" (2008, p. 34). Instead of isomorphism, the push on this account is for complementarity between brain and world.

Such complementarity furthermore takes place among a nested hierarchy of possibilities. Following Shepard (1984), we can consider the case of how a piano resonates with external sounds to better grasp how this nested hierarchy works. If one is standing near a piano with a guitar and plucks the C-string, the piano will start resonating from the soundwaves in the air. At the same time, it is not only the piano's C strings that vibrate in response to the original guitar note. Other strings are excited and vibrate too. This result is further augmented in the case of complex inputs. There can be a variety of "modes of resonance" that occur in response to complex stimuli, such as an entire chord being strummed instead of an individual note, and sometimes resonance for an organism can even occur without standard environmental inputs, such as during dreams or hallucinations. The level of activity involved with these different modes also varies in important ways that require further consideration.

The second option comes from Vicente Raja's (2018) proposal to combine Michael Anderson's work on neural reuse with Dynamical Systems Theory to establish an anti-representationalist account of resonance. Neural reuse follows a similar set of principles to

resonance in general, wherein different areas of the brain will resonate together in response to the demands of the task at hand. The additional dynamic coupling between these two – neural reuse, on one hand, and the environmental scale, on the other – is held together in relation to a common ecological variable. In doing so, we are able to say that the dynamic intra-individual system (i.e. the agent) *resonates* with the dynamic environmental system.

Raja further postulates that the resonance between these two dynamical systems is primarily a one-way phenomenon, wherein the ecological scale constrains the intraorganism scales but not vice-versa. In his words, "*to explain resonance is to account for the coupling of the dynamic systems at the ecological and intra-organismic scales in terms of the ecological variable that constrains a given agent-environment interaction*" (2018, p. 41, emphasis in original). While this overall approach offers important additional insights, many of which I have not canvased for the sake of space, I am nevertheless concerned with this part of his final claim.

To note more about my concern first, however, I am not concerned with the idea that the environmental constraint may restrict an agent's behavior in a manner outside of their control. Nor am I concerned that the focus on an ecological variable would return us to the ecological level without explaining the brain as resonant organ. Instead, I suggest that the problem comes from placing the environmental variable as one to which an organism is necessarily constrained without accounting for the fact that an organism can, in many cases, alter these environmental pressures. Such alterations can take place both in the moment and over the course of a lifetime (or multiple lifetimes, if we are considering a species and not just individuals). Raja's account as it stands therefore does not yet explain the role of how organisms actively modulates their niches and the rest of the environment in real time. These sorts of modulations are especially integral parts of musical performance and the development of different musical genres.

Likewise, the question of novelty returns as a problem here as well. While perception in jazz performance requires similar sorts of matching as perception in other contexts – in order to perform, an improvising musician must be attuned to what has been played and is currently being played, just like someone walking down the street must be attuned to what is happening at their feet and what is happening around them – the jazz case goes beyond it as well. Jazz improvisers place a premium on novelty and unique engagement with the environment, regardless of whether they stay close or move far away from source material during performance. While novelty and uniqueness may be important for everyday engagements with the world, the general role of perception is not to come up with unique interpretations of the environment, but to properly orient us towards it. The same may be true for part of improvisation concerning the fact that perception is needed while improvising. But such close orientation to the world does not seem to be true for all of improvisation. Insofar as Raja's account does not account for this latter case, we either need to turn to Nussbaum's account or provide additional ways for developing Raja's account moving forward. I take both of those options as live possibilities for future development.

Chapter Conclusion

In this chapter, following Risjord's ecological theory of action, I considered additional ways that attunement to affordances mediated by meta-cognition can be the bedrock for explaining action and agency. I further developed these three areas by using Coleman Hawkins' performance of *Body and Soul* as a primary case study, along with additional examples as needed. While I reached a series of responses to key aspects of attunement to ecological information and affordances vis-à-vis their metaphysical underpinnings, the last section on meta-cognition did not conclude with a clear stand on the best account of resonance. What more should we add to the

idea that the brain resonates to explain the functioning of meta-cognitive capacities in this case? One option is to consider it in light of predictive processing, as I do in the next chapter.

While it is important for a successful musical performance that a musician matches the environmental patterns of the song, they must also go beyond matching towards novelty and creative engagement with environmental patterns. The brain, in other words, is not just an organ of resonance for perceiving the world, but a creative organ that must actively act on and within the world. It is likely, given the speed at which musical improvisation takes place, that some of these actions will need to be planned out, anticipated, or at least predicted ahead of time. We would be dull, boring creatures if the only things that a brain could do is help us better match what is currently going on in the world, and we would be clumsy, most-likely-extinct creatures if all we could do is respond to the happenings of the environment just a bit after they happen.

Thus far we have been accepting a central assumption of non-computational approaches that are integral to orthodox ecological psychology, along with exploring some ways in which the concept of resonance may account for the role of the brain in the complex brain-body-environment system of cognition. However, though certain parts of anticipation have been explored so far – for instance, affordances imply some kind of view towards the future and possible action as much as the realization of particular concrete actions – we must now turn towards a more detailed account of the anticipation process. I propose to do so through a rather recent view of the brain as a "predictive machine." How these predictions are best captured, especially if they require some sort of representational capacities above and beyond resonance with environmental information, will be a core topic of the next chapter.

Chapter Four: Improvisation, Surprise, and Predictive Processing

The world is a surprising place and an essential part of improvisation helps us to navigate surprises. Walking into a college classroom with an elephant, for instance, would normally be more surprising than being cut-off in city traffic. Yet surprise also comes in degrees; the traffic situation would be more surprising if the offending car wasn't seen until the last possible second.

In order to improvise and deal with surprise, an agent must project ahead in time. This projection includes anticipation and prediction in regards to one's own actions as well as the actions of others. The need to anticipate about others is particularly acute for musicians and performing artists, where ensemble members face strong and pressing time constraints. Prediction and anticipation are thus tied to the present as well as the future. They also draw on abilities to remember what has happened in the past too.

The idea that music performance and music perception involve predictions of various kinds is not a new claim. It has been widely acknowledged in the literature on music cognition, and especially in areas such as the relationship between music and emotions as well as music and performance (Huron, 2006; Schaefer, 2014; Leman, 2016; Keller, Novembre, & Loehr, 2016; Meyers, 1956). Predictions and anticipation serve to maintain a stable performance environment, especially when performing at fast tempos and/or coordinating within an ensemble. Another role for predictions in music comes from how a song fulfils or frustrates a listener's expectations, especially regarding genre norms. How to best characterize these sorts of predictions and anticipations remains a live question that requires further exploration and addressing this issue is a main focus of this current chapter.

One recent and useful hypothesis for exploring the nature of anticipation in improvisation comes from Predictive Processing (PP) accounts of cognition (Clark 2013; 2015; 2016), also

known as Predictive Coding (Hohwy, 2013) or the Bayesian Brain Hypothesis (Friston, 2012). PP, if correct, offers an extremely strong explanatory tools for understanding minds. It is further intriguing in this regard for multiple reasons, especially its purported offering of an overarching theory of cognition around the idea that an agent's primary goal is to decrease surprise and better predict future inputs. Both of these goals may also relate to an organism attempting to maintain a better grip on the environment. (Rietveld & Kiverstein, 2014)

While PP is in general focused primarily on the brain, it offers a possible embodied account as well. As Andy Clark (2016) has noted to the idea that the brain is in the business of making predictions about future inputs:

> we must locate the inner prediction engine in its proper home. That home…is a mobile embodied agent located in multiple empowering webs of material and social structure. To make full and satisfying contact with the thinking and reasoning of agents *like us*, we must factor in the myriad effects of the complex social and physical 'designer environments' in which we learn, act, and reason. (p. xvi, emphasis in original).

One of the most potent designer environments is music, as we shall return to again below. As a purported overarching principle of the mind and cognitive activity, however, there is an important question as to whether or not PP could synthesize and incorporate the insights from ecological psychology, especially those highlighted in the previous chapter, with other sources of research. As of now, there is no consensus on the answer to this particular question.

The exact scope of PP is an open question. First, as noted above, PP has been pitched as an organizational principle for the brain and perception (Clark, 2013). Second, PP, at least in regards to the role of the Free Energy Principle (FEP), has been pitched as an organizing

principle of life itself (Friston, 2013). Neither of these approaches is mutually exclusive, and recent accounts of PP often bring the two together (Ramstead, Veissière, & Kirmayer, 2016; Ramstead, Kirchoff, Constant, & Friston, 2019; Linson, Clark, Ramamoorthy, & Friston, 2018). Furthermore, some accounts of PP take the sole organizational principle of the brain to be Bayesian[1] in nature, while others take it to be a primary functional organization for cognition without being a totalizing mechanism to explain everything (Drayson, 2017).

An additional feature of PP is that the brain constantly takes part in a fundamentally active process. Such activity is, first and foremost, geared towards cognition and action *in the world*, rather than a more passive cognition or cognition that is geared solely towards absorbing and processing abstract information. Moreover, the brain is active as a *predictive* machine, constantly predicting the input that will face the system in the future as well as continuously adjusting future predictions in light of current feedback from the world.

Some of these predictive processes are offloaded onto other parts of the organism, as in the case of the structure of the retina (Srinivasan, Laughlin, & Dubs, 1982; Johnston et al., 2019), which develops to start actively filtering input based on previous experience to the agent and the visual system, even at the first entry of visual information into the retina itself. Another prediction offload is the anticipatory nature of eye saccades (Vetter, Edwards, & Muckli, 2012). The sorts of adjustments made in the brain and body, as well as in the brain-body system working as an inseparable unity, can therefore operate at multiple levels and timescales, and they manifest on a range from the ability for an outfielder to quickly adjust their movements in order to catch an incoming flyball in seconds to a musician learning how to play an instrument across a lifetime of practice, and many different scenarios and timescales in-between.

[1] An account is Bayesian if it follows Bayes' theorem. More information on this theorem are considered below on p. 169-171.

In addition to addressing these issues, I shall focus on questions of novelty in PP and music performance. While this focus on novelty is particularly important in jazz for aesthetic and pragmatic reasons, part of the link between novelty, PP, and performance requires considering further the case of music and surprise. Then, after exploring the issue of novelty in general, I turn to consider a possible framework that synthesize elements of shared intentions, affordances, and predictive processing.

In addition to the framework, the role of novelty, and the relationship between ecological psychology and PP, there are also important debates concerning whether or not PP is a friend or foe to embodied cognitive science. According to Jakob Hohwy (2013), these research areas are incompatible and accepting PP requires returning back to a disembodied, internalist picture of old. While Hohwy takes this conclusion as a fine result, it would not be so fine for me, since my goal remains exploring how *embodied* cognitive science can be relevant to understanding jazz improvisation.

In contrast, according to Clark (2016), PP and embodied cognitive science are compatible. Clark further argues that PP offers the most plausible account of the role that the brain plays for embodied accounts. Finally, on even more radical embodied versions, such as the enactivist one provided by Shaun Gallagher and Micah Allen (2018) or another enactivist version proposed by Dan Hutto and Eric Myin (2017; see also Hutto, 2018), PP and embodiment are in principle reconcilable, but this reconciliation process will require a major reconceptualization of core PP commitments to jettison any talk of concepts such as 'representations,' 'inference,' or 'internal models.'

Pace Hohwy, along with following Clark and the enactivists, I suggest that there is a strong link between PP and embodied approaches. One remaining concern here is whether PP, at

least as so far conceived, is inherently a cognitivist endeavor. Hutto and Myin (2017) are particularly vociferous about this point. As a result, for now I will simply say PP is embodied since the implementation details of its realization in action fundamentally matters. For example, affect, as a fully embodied phenomenon, is thoroughly interwoven into the updating and deployment of predictions, and we cannot understand PP without placing a role for embodied affects in it (see Barrett & Bar, 2009). Our predictions, in turn, are important parts of what and how we see the world. Taken together, there simply cannot be any neat input-cognition-output distinction once PP is part of the cognitive story, and thus at least one of the main tenants of disembodied cognitive science is at odds with the PP approach.

While we can tentatively reject the strong cognitivist reading of PP, there is an important caveat to be added in regards to levels of explanation in this case. I accept that, if our target level of description is the neural level, then the nested cascade of predictions in the brain can be understood as not being in direct contact with the external world. One reason for this conclusion is a potential category mistake if we hold the opposite position; it is, after all, an agent that sees and interacts with the world, not a bundle of neurons in isolation.

There is also the issue of whether or not the predictions of PP require additional mediation in the form of representations or not. I will not provide an argument in favor or against representations here. Instead, I wish to note that, insofar as the nested neural models of PP are tuned to gain an optimal grip on the world, they are not aimed at re-presenting it in perception nor are they aimed at mediating access between organism and environment. Instead, they are aimed at linking up the organism to the world as best as possible. Whether or not this linking process also requires (neural) representations is an additional question.

Perhaps the best way to further bolster the directness of PP at the agent level, again *pace*

Hohwy, is through Karl Friston's claims that (1) the agent is a model of their environment and (2) an environmental niche is a model of an agent. As noted in the previous paragraph, the agent is not just the brain of an organism. As an agent learns about various parts of the world, their predictions about future sensory sensations will become more accurate over time, along with learning how to better capture their needs, desires, and the sensory situations they are likely to face in day-to-day contexts.

This perceptual accuracy is built, in part, from repeatedly sampling features of the world, along with taking part in repeated interactions in the environment at large. As we saw briefly with Clark's concept of "designer environments," the environments that we learn to better predict are simultaneously the products of our collective human and non-human animal activities. In the case of humans, these designer environments are fundamentally socially constituted. Learning about how to engage with them involves enculturation into social and environmental niches. A designer environment furthermore provides support and a place for supervised feedback during learning for infants, children, and adults alike that further shapes and sculpts the nature of their predictions (and the reduction of surprise about the world).

While I will not be taking a definitive position among the kinds of embodied accounts that could be linked with PP, I still hold that a strongly internalist, disembodied understanding of it is doomed to fail. Part of this failure is for reasons noted above. An additional concern is that an internalist account of PP will fail to properly account for flexible cognitive boundaries that are part and parcel of prediction error minimization.

To see why in more detail, let's consider the temporal flow of conscious perception. Jakob Hohwy, Bryan Paton and Colin Palmer (2015) take an internalist stance towards temporal flow and argue that our phenomenological experience of time is driven by an in-built mistrust of

the present. While I agree that there is an important role for the guiding assumption that the environment will not stay static over extended periods of time, this position does not account for the ways in which an individual agent can selectively alter higher-order assumptions about the flow of error itself. Some of this higher-order manipulation may be direct, such as counting the seconds as they pass to maintain a steady trust among an otherwise changing environment. Others may be more indirect, such as trying to put oneself into a better or worse mood to control the feelings associated with changes in time. This selective alteration itself could include elements of bodily processes as well as "active inference." In other words, placing a premium on mistrust of the present may too strongly discount the fact that an agent may have some conscious control over how much they do (not) trust the present. Distrusting the present likewise does not account the ways that the environment can positively solicit actions from the agent, which is an important part of a well-executed musical performance. Finally, the strong internalist does not, and simply cannot, account for the fact that PP is meant to include various features of the body and environment that are not limited to neural activity alone.

While the above may help establish why we ought to favor an embodied version of PP, I have yet to say much about what the main alternative to perception (and action) is or could be yet. As a result, the next section will be focused on providing a deeper overview of PP as well as its main alternative in the feature detection model.

Feature Detection vs. Predictive Processing: Two Conflicting Accounts

To clarify the grounds for PP, it is helpful to contrast it with the heretofore dominant account of perception as feature detection. According to proponents of the feature detection model (Martin, 1994), perception occurs when an organism builds up increasingly complex representations of the world as information travels through their sensory systems. The build-up of perception on

this account, in addition to being computational, is largely a passive process. The passivity comes from the fact that the flow of information moving through the brain, slowly builds up in complexity as it is passed along from more basic levels to more abstract levels of the perceptual hierarchy, and results in a percept. Likewise, while attention has been shown to play an important role in feature detection models, the main focus has not been placed on active attention.

For a concrete example, consider the case of seeing a puppy. On the feature detection model, perception begins when various environmental stimuli reach the sensory systems; light hits the eye, sound hits the ear, smells hit the noise, etc. The final percept – a puppy, in this case – is derived from the information gleaned through these inputs. Early parts of the visual system, which have been found to only respond to detecting light and specific static bars in part of the visual field, begin the start of the process. Later parts of the visual system are sensitive to movement and increasingly complex kinds of properties. Both lower and higher-level features are eventually modulated at some point by top-down conceptual and general-purpose information. When both top-down and bottom-up work in concert for seeing the puppy, not only do we see a coherent visual object of a puppy, but we see it as a *dog* instead of as a *cat*.

The feature-detection model has been subject to some well-known criticisms. One major issue is the binding problem, which concerns how inputs from different sense modalities (such as a bark in audition, the smell of wet fur in olfaction, and the sight of a yellow blotch in vision), or how inputs within a particular sense modality (such as a mass of yellow fur, a black nose, and red tongue against the background of brown tree trunks and green grass), come together as one coherent object instead of existing as several unrelated stimuli. A common suggestion for solving this problem in intermodal cases comes from temporal similarity (Treisman, 1999), such that the

bark and the blotch of yellow fur, when they appear close enough in time, are taken to be coming from the same object. Other non-temporal cues may help distinguish among different objects within the same sense modality as well as distinguish objects and the background. Figure-ground distinctions derived from tracking feature discontinuities, for instance, has been offered as one possible explanation for how we can see multiple objects in the same environment, at the same time, instead of one undifferentiated object (Poort, Raudies, Wannig, Lamme, & Neumann, 2012). In addition to harboring these assumptions, however, the feature detection model offers a disembodied approach to perception.

This link between disembodiment and feature detection models may be made clearer when we consider the case of music. Eric Clarke (2005, p. 11-16) offers a succinct overview on how the feature detection approach has been applied to music cognition. Starting with raw auditory input of soundwaves bouncing around in a medium, the auditory system extracts increasingly complex mental representations, moving from psychoacoustic properties such as pitch, timbre, and auditory streaming, to more general cognitive properties such as form, tonality, and scale, then finally culminating with still more abstract and general features such as aesthetic value, reference, and musical meaning,

While proponents of this account note that the bottom-up flow of increasingly complex information is attenuated by some top-down features, there remains a heavy amount of cognitive lifting on the side of the organism to make sense of sensory information as it moves from sensation to cognition and cognition to action. This weight becomes clearer in cases of more abstract musical meaning, wherein different musical meanings supervene on the same psychoacoustic features and some musical meanings can only be known with certain cultural background knowledge. In these sorts of cases, even an exhaustive account of possible musical

meanings does not exhaust what a listener may hear at a given time, or to what features a musician attends to during performance. Someone who knows taps will be able to hear it as part of Jimi Hendrix's famous performance of the Star-Spangled Banner at Woodstock, while those who are not may hear something else or nothing specific at all (see also Clarke 2005, Chap 2).

The PP alternative to feature detection accounts does not deny that there are flows of bottom-up and top-down information in perception. Instead, it fundamentally reorients our understanding of the role these information flows play in cognition. To begin, while "top-down" and "bottom-up" are often invoked in the context of psychological explanation (Engel, Fries, & Singer, 2001), their exact meanings have become muddled over time. For instance, Karsten Rauss and Gilles Pourtois (2013, p. 2), following Andrew Engel, Pascal Fries, and Wolf Singer (2001), note that there are at least four different ways of understanding top-down processing: anatomical (functional activity among descending levels of the hierarchy) , cognitivist (an explicitly hypothesis-driven understanding of top-down processing), gestaltist (top-down processes provides contextual modulation on bottom-up processes), and dynamicist (higher level entrainment of local neurons via oscillation with others, sometimes widely distributed across brain regions).

Depending on the account of PP in question, different understandings of top-down are treated or less as more important. For instance, an enactivist version of PP would likely foreground an anatomical and dynamicist reading and downplay or even reject a cognitivist understanding of the phenomenon. Clark's version, in contrast, may take elements from all four of these different senses of top-down. For Rauss and Pourtois, top-down processing occurs in cases where one or more levels of the anatomical hierarchy are skipped, allowing for rapid modulation of lower-level processes that are normally fairly rigid in their functioning. In their

words, for a connection to be top-down on PP "there must be a direct anatomical connection between a source and a target region, and the source must be located at least two levels above the target" (2013, p. 6).

Rauss and Pourtois furthermore highlight two additional points about the top-down/bottom-up distinction in regards to PP that are worth noting here. First, we should *not* understand top-down and bottom-up as opposed processes where they compete against each other. Instead, they work together to form a cooperative information exchange with the mutually supporting aim of minimizing prediction error in the system. Second, the bottom-up/top-down distinction crosscuts various sorts of ascending and descending connections, and thus operates laterally – e.g. there are connections between various neurons in V1 - as well as vertically – e.g. there are connections between various neurons in V1 with various neurons in V2 - in the system. These connections serve to either inhibit or excite additional processing, which in turn allows for a combination of rigid, slow processing and fluid, fast processing. Combining these two elements maximizes the ability for an agent to successfully respond to the environment.

In short, for PP we find that the top-down information flow is best understood as a cascade of vertical and lateral predictions along a nested hierarchy of levels within the brain (and also, for some instances, across levels in the body and/or in the world). Predictions are also part and parcel of generative models. These models capture the combination of (1) a conditional probability about a certain state of affairs and (2) the observed state of affairs at a given time. In each level, the system is constantly in the process of predicting the inputs that are flowing up from the level immediately below. If there is a conflict between the actual input and the predicted input, information is passed up to the next higher level, which has its own constraints and predictions.

The bottom-up information flow in PP is not accrued into a more complex representation or percept as it moves up levels of perceptual processing. It is, in contrast, simply not in the business of carrying along any increasingly complex representations. It rather serves to pass along an error signal that indicates a mismatch between predictions and actual sensory input. If there is a match between prediction and sensory flow (i.e., in cases where there is no prediction error), no further bottom-up information is passed along. We can therefore understand bottom-up information as a check against the predictive model produced from the level immediately above.

Another essential component of PP is the free energy principle (Friston, 2013). This principle is tied into the process of free energy minimization (FEM), which is similar to prediction error minimization as discussed above. In thermodynamics, free energy is a name for the energy available for an entity to do work. In information-theoretic terms, free energy concerns the discrepancy between the average, expected states that an organism faces and those that it is actually experiencing at a given moment. The technical link between these two is reflected in the underlying mathematics, but again can be easily understood as the mismatch involved in prediction error. There is also an additional link between the two senses of the term insofar as too much free energy in either the thermodynamic or information-theoretic senses is bad for an organism. At the extreme end, if an organism is unable to keep its free energy within a certain acceptable set of boundaries (i.e., if an organism is unable to successfully minimize its prediction errors), it will die.

Excess amounts of free energy can be correlated with surprise. While some surprise may be conscious – consider, for example, the different sorts of surprise when opening a door to find a distant friend standing on the porch vs. an aggressive and hungry tiger pacing around – equally important cases are not. To distinguish between them, a subpersonal notion of surprise, called

"surprisal," has been introduced (Clark, 2013). Surprisal occurs whenever there is an unexpected variation from the general statistical patterns that are expected to hold in a situation.

Both surprise and surprisal are therefore related to situations that veer from what is expected, but the relevant kinds of subpersonal surprisal that we are interested in will vary depending on which level of the predictive hierarchy we are concerned with at a time. Low level parts of the visual cortex, such as V1 and V2, may be focused on predictions of edges and basic movement, while higher levels may focus on faces, and still higher levels on certain patterns of social activity that are associated with standing face-to-face with the Queen of England.

Surprisal is a variable that cannot always be controlled directly by an organism. In order to address dangerous build-ups of it and attendant free energy, there are two main options. First, an organism could update their predictions, which would help them better grasp inputs by making more accurate guesses about the world. Unfortunately, although this is an essential process, updating predictions alone does not serve to decrease free energy directly. Part of why comes from the fact that changing predictions can't capture distinct needs for how organisms are embodied. Having a body entails additional conditions that must be met to survive besides knowing what is happening in the environment (see also Bruineberg, Kiverstein, & Rietveld, 2018). Even if I am able to make perfect predictions about the perceptions associated with being adrift in space without a space suit, doing so while floating around will not decrease my free energy nor will it increase my chances of survival.

The second alternative is for an organism to keep its predictions the same while moving about so the world more accurately matches the predictions. This option is more formally known as *active inference*. Unlike changing predictions to fit inputs, active inference allows an organism to directly decrease available free energy. In the spacesuit example, active inference

would call for putting on a spacesuit (or at least trying to put one on if it's not already too late) to get back to the expected state of being able to breath, among other things, instead of coming to terms with the fact that one is in space without a spacesuit.

Since an organism learns over time how certain patterns of prediction and action are related, they will eventually be able to predict outcomes in advance (enactivists call this sensory-motor contingency; e.g., Nöe (2004)). As a result, proponents of PP have noted that the same process for decreasing surprisal in perception can be flipped to drive endogenous action. On the PP account, action occurs when a system introduces predictions that preempt the expected action while simultaneously inhibiting the use of prediction-changing to mitigate the surprisal now present in the system. If a person wants to do something such as raise their right hand, they would begin by predicting that their right hand is already raised. Making this prediction could be done either consciously or unconsciously but, in either case, since the current sensory state ("both of my hands are on the table") does not match the predicted state ("my right hand is raised"), there will be an increase of error in the system between prediction and reality. By utilizing active inference, the way to decrease the prediction error would be to raise my right hand in line with the original goal of creating the prediction error in the first place.

For an additional example of free energy minimization in action, consider the case of a trained listener told that they will be hearing a musician play a C scale. While listening to the scale, they will not only hear the immediate first note being performed. They also anticipate, and thus predict, that the next note to be played will be a D. If the next note is, in fact, a D, then there is no need to update of the system since the predictions are correct. If the next note is not a D, then prediction error would be generated and the system must revise its prediction.

Hearing a note, such as C or D, on the feature detection model starts with hearing a

particular pitch, passed along the auditory processing hierarchy, and ultimately categorized as one of the respective notes, as well as the correct first and second notes of the C scale. However, for PP, the trained listener is already predicting that a D will happen the moment they know that a C scale is being played, since they have a relevant set of expectations about what a C scale is and how it progresses from note to note based on prior experience, prior knowledge, and/or prior beliefs; often these various cases are simply grouped under the term "priors." Such priors are furthermore an important part of generating various predictions and are updated according to relevant feedback they receive over time.

In hearing the musician play a D# instead of a D, no amount of moving around the world would fix the mismatch for a listener. They would thus either change their prediction ("perhaps this is a different scale and I was told the wrong thing/mistaken at first") or, in this case, grant that the musician is simply wrong to play that particular note. The musician, in contrast, and assuming they were planning on playing a C scale, would be able to take actions to address their incorrect note, although that may only be by stopping, tacitly or explicitly admitting a mistake was made, and starting over again. If we expand the context from playing a scale to playing a bebop solo, the listener would remain in the same situation as before. In contrast, the musician, in realizing they played the 'incorrect' note during a solo, will attempt to address it as much as possible and as required within the context at hand.

While there are times where only one path to prediction error minimization is possible for the brain to use, such as the cases charted in the previous paragraph, a combination of revising predictions and active inference in tandem is important to address surprisal in many cases, such as learning to play an instrument or becoming able to understand an unfamiliar style of music. Utilizing both techniques is indeed a more common alternative for dealing with everyday tasks

and even more so common when we consider that the brain is not an isolated part of the organism, but an essential partner in the dance of brain, body, and world.

Thus far we have been considering the nature of predictive processing more or less on its own merits. However, recent developments have also integrated it into larger frameworks. Exploring these frameworks is essential for at least two reasons. First, they allow for further addressing how PP can be taken as a full part of the brain-body-world *explanans* that I am considering through embodied cognition. Second, frameworks allow us to bring together the insights from Chapters 2 and 3 concerning joint action and ecological psychology to present a unified model of bebop performance, which I will consider in more detail in the final conclusion.

There are at least three existent frameworks available for this task: the Active Inference Framework (AIF) (Linson, Clark, Ramamoorthy, & Friston, 2018); the Skilled Intentionality Framework (SIF), (Rietveld, Denys, & Van Westen, 2018); and the Cultural Affordances Framework (CAF) (Ramstead, Veissière, & Kirmayer, 2016; Veissière, Constant, Ramstead, Friston, & Kirmayer, Forthcoming), All three start with the assumption that there is a congruence between embodied approaches and PP, although the nature of this relationship varies in important ways between them. In what immediately follows, I will focus on the AIF, then return to consider the CAF in a later section. The SIF is not considered, first, since it was explored in some detail in the previous chapter and, second, insofar as CAF is an explicit extension of the SIF, they have substantive theoretical overlap. Similar points can thus be made for both frameworks, *mutatis mutandis*.

Before turning to these frameworks, however, a few words must be said on bridging PP and ecological psychology in the first place. In addition to what was considered in the previous chapter, another serious concern raised by Shaun Gallagher (2017) is that a weak version of

embodied cognition, involving brain-based representations and found in certain versions of PP, cannot be combined with radical versions of embodied cognition, especially of the enactivist variety. I agree with Gallagher that these two versions of the 4E paradigms would not combine well within one hybrid model. However, I propose that both of the two main alternatives between Clark's version of PP and the enactivist version escape this hybrid approach by either sticking more closely to enactivism or towards a slightly more conservative version of embodied cognition. By exploring the AIF (which more closely embraces Clark's approach) I will now introduce several additional general assumptions underlying PP. Later, by considering the CAF (which draws heavily on enactivism) I will more directly address the link between PP and ecological psychology.

The Active Inference Framework (AIF) takes the free energy principle and FEM as foundational principles. Moreover, an important addition to the FEM process considered above is the fact that the AIF explicitly focuses on the rich interplay between exteroception, interoception, and proprioception. The fusion of these three different types of perception are an essential feature of music performance, which inherently involves a tight dance between self and other made sounds in the environment, the metacognitive tracking of those sounds, interoceptive feelings around what is being played, and a command of the proprioceptive movements that are needed to create and maintain sounds throughout performance.

Adam Linson, Andy Clark, Subramanian Ramamoorthy, and Karl Friston (2018) further offer a redescription of a Gibsonian approach that meshes with their understanding of PP. Starting from the idea that generative models are best understood as part of a model space (2018, p. 10), they note that the environmental action opportunities at the heart of ecological psychology are best understood as exteroceptive models. These models, when paired with affordances, are

relevant to action perception within the overall model space, but they are not the whole story. The building of the model space through multiple types of overlapping predictions means that certain higher-level invariant correlations are developed and utilized when needed.

One multi-purpose example of the above points can be seen in attempting to cross rapids by using as series of stepping stones (Linson et al. 2018, p. 13). While rapids are constantly in motion and flux, the stepping stones appear as "local minima of uncertainty." In a literal reading, the stepping stones are examples of affordances that allow someone to cross an otherwise uncrossable river. In a metaphorical reading, the stepping stones are reliable invariants among the thermodynamic flux around them that allows an organism to gain a grip on its environment, act, and, achieve its goal.

Affordances are furthermore taken on the AIF under the guise of policies for action, which could be related back to previous work canvased in Chapter 2 on Bratman's notion of planning agency and policies that underlie it. For both Bratman and proponents of the AIF, policies are related with plans for action. However, the AIF policies are not necessarily the action plans that were prominent in early robotics approaches or in the sense of full-blown folk plans. Instead, policies are concerned with transitioning between the current state of affairs to a preferred future state, in turn utilizing the combination of interoception, exteroception, and proprioception to drive action. They thus can be related to the earlier explanation of action production on the PP model.

I have thus far been considering the core features of PP with minimal references to music performance. Before turning to directly combining PP and bebop improvisation, the following section will address this lack by taking an important detour into research on music, surprise, prediction, and expectation. Doing so will help set the scene for considering PP and novelty, as

well as introducing the CAF as an account that can be used to address both novelty and the insights raised by these existent accounts of music and surprise.

Music and Surprise

Leonard Meyer, in his classic text *Emotion and Meaning in Music* (1956), argues that musical meaning is a product of embodied expectations (Meyer, 1956, p. 35). These kinds of expectations can be either conscious or unconscious. In either case, they are tied to learned sets of expectations that are developed as a listener engages with a particular style of music. In the case of bebop, for instance, learning to play a 'successful' solo is in part about properly following the musical possibilities allowed within the preset chord changes of a tune. Since bebop practices varied from earlier forms of jazz improvisation, it originally received pushback from earlier jazz performers and listeners alike as meaningless music. Older performers and listeners had yet to develop the skills to appreciate bebop music in its own rights, in addition to holding onto other assumptions they had about what counts as "good" music. Even within traditional bebop constraints of following particular chord changes while soloing, however, there is still a wide swath of possible "acceptable" paths of what can be played. Sometimes an expressive solo with many wrong notes may be preferred over all the right notes without soul.

Musical expectations can be various and nested. The way in which they are nested will ultimately depend on a variety of considerations, including what features are in line with a person's habitual modes of listening. In Meyer's words about how musical meaning comes into focus, we find that "given a mind that is disposed towards objectification [i.e. when self-consciously focusing on the music raises it to a consideration as an objective thing in consciousness], meaning will become the focus of attention, an object of conscious consideration, when a tendency or habit reaction is delayed or inhibited" (1956, p. 39). Such

tendencies and habits can be changed over time, of course, so exactly how they are delayed or inhibited will change as well. Different affects correspond with different ways of delay or inhibition, and different genres will require learning new habits and tendencies for performing and listening practices.

Meyers work has served as an inspiration for future work in the field. One of the most comprehensive studies on the relationship between music and surprise to date was developed by David Huron in his book *Sweet Anticipation* (2006). According to Huron (2006):

> Despite Meyer's interest in psychology, however, *Emotion and Meaning in Music* was written at a time when there was little pertinent psychological research to draw on. In the intervening decades, a considerable volume of experimental and theoretical knowledge has accumulated. This research provides an opportunity to revisit Meyer's topic and to recast the discussion in light of contemporary findings. The principle purpose of this book is to fill in the details and describe a comprehensive theory of expectation—a theory I have dubbed the 'ITPRA' theory. (p. 2-3).

The ITPRA theory, while primarily applied to music by Huron, is meant to capture expectation in general. The acronym refers to several anticipatory mechanisms that form the basis for expectations and the relationship between expectation and emotion: Imagination, Tension, Prediction, Reaction, and Appraisal (hence 'ITPRA'). Focusing on the case of music, this model can be applied to understand expectation in listeners and performers. Having application to both of these groups is a benefit of Huron's work and makes sense from an embodied perspective, wherein even apparently passive listening to music is an active process (Moran, 2014).

One of the main differences among these mechanisms is whether or not they occur before

and after event onset. According to Huron, Imagination and Tension are pre-event onset mechanisms, while Prediction, [2] Reaction, and Appraisal are post-event onset mechanisms (Huron, 2006, p. 8-15). To see these five kinds of anticipation in action, we can consider the case of a musician about to play *All the Things You Are*. For simplicity sake, let's focus on the moments leading up to a sax player beginning the song and, even more narrowly, just on the first two notes of the tune.

Before the song begins, the saxophonist has a sense of what they will play, how they will play, and when they will play the first two notes. This sense is captured by predictions about the what, when, and how of these first two notes. Some of the motivation for playing both in general and for playing these specific notes in this way will be provided by the mechanism of *Imagination*. Especially for an expert musician, imagination is not just about abstractly capturing notes and theory. It is laden with affective expectations about the quality of the sounds they plan on producing and how each note should relate to what is being played by the rest of the ensemble. These moments of imagination will help bring joy to the expert and maintain the novice as they work through the highs and lows of learning their instruments.

Just before the song begins, the musician will experience increasing *Tension*. Some parts of tension have a phenomenological component, like the feeling of butterflies in the stomach before a performance. While these phenomenological feels are relevant, the mechanism of Tension is concerned first and foremost with the embodied preparations required for action, rather than specific feelings of tension. For instance, an increased heart rate, perspiration,

[2] Since prediction implies something happening before event onset, one may rightly ask how is prediction a post-event mechanism. To clarify, we must distinguish between the original prediction and whether or not a prediction was fulfilled by the situation. The original predictions that feed into the Prediction mechanism are developed pre-event onset. However, the Prediction mechanism functions by tracking whether or not these predictions are supported during the event. As a result, it is considered a post-event onset mechanism since the event onset needs to occur before predictions are fulfilled or frustrated.

changed breathing, pupil dilation, and purging as many distracting thoughts as possible from consciousness are several changes that occur as a result of Tension. It in turn helps facilitate motoric and perceptual preparation even before the song begins.

At the moment of event onset when the first, then second, notes are played, there is an immediate matching (or mismatch) concerning the *Predictions* of how the performance was expected to unfold. Matching prediction to what actually occurs will result in a positively valanced response, and negative responses will result from mismatches between predictions and reality. These predictions are further paired with either a fast or slow response mechanism. The immediate fast *Reaction* to any (mis)matches operates outside of consciousness and has a fast onset time. It also often assumes a "worst-case scenario" As the performance unfolds, the musician is able to *Appraise* certain parts of it. Sometimes appraisal and the reaction are mutually reinforcing, while other times they are in disagreement. Assuming that the performance is recorded, they are also able to go back and relisten to what it for further Appraisal.

At first glance, since expectations and predictions are central features of both ITPRA and PP, it seems that these two can be integrated into one overarching model without much difficulty. There is indeed nothing about either that automatically forecloses their union into one model. At the same time, while expectation is the core of both the ITPRA model and PP approaches, just having predictions alone is not enough to make an account or model a variation on PP. In addition, as captured by the understanding of PP as the "Bayesian Brain Hypothesis," there is an important historical link between PP and a Bayesian approach.

Nico Orlandi (2018) has argued that there are important differences between whether an account is predictive in nature and whether an account is Bayesian in nature.[3] In order for a

[3] She further argues that it is possible to have non-Bayesian PP accounts, which I do not consider in more detail here solely for the sake of space.

system to be Bayesian, it must at least approximate Bayes' theorem, which entails a specific set of constraints on how an organism updates its prior beliefs in response to environmental change. Updating beliefs in line with Bayes' theorem will require altering precision weights of priors and making adjustments to how much trust the agent places on predictions instead of current sensory input, or vice-versa. More concretely, one of the main Bayesian constraints entails that an organism must shift all their prior credences – that is, they must change the amount of weight they put on one prediction or model over all alternatives - in light of the current posterior possibility among these different alternatives.

While the shifts of prior credences are often considered in terms of adjusting precision, being a properly Bayesian agent further requires changing all beliefs in light of posterior probabilities. The reason for changing all beliefs stems from the assumption that Bayesian agents are optimally rational. To be optimally rational, in statistical terms, the base possibilities among all potential options that result in an event must equal 1. Thus, if one option becomes more likely – say moving from .58 to .75 – the other options must have a decreased weighting. Otherwise, if alternative possibilities are not adjusted accordingly, the agent is not acting as an optimally rational agent on Bayes' terms.

To see what these technical terms mean in more ordinary usage, consider the case of a person finding a wet patch of grass and deciding between two possible cues of what caused it to become wet: (1) a sprinkler system or (2) rain.[4] In this case, their choice will be a combination of prior evidence, i.e. priors, about the likelihood of whether the wet grass was caused by (1) or (2), on one hand, and "conditional credences," such as whether or not the sun has been out all day or if the sprinkler system just finished operating, on the other.

[4] This example and its subsequent modification are taken from Orlandi (2018, p. 2376-7).

Each of these credences are granted a weight between 0 and 1, and their sum will be a total of 1. Moreover, following Bayes' rule requires operating over all relevant beliefs and predictions, as well as operating "by exoneration" and "backward blocking" (Orlandi, 2018, p. 2376-7). Therefore, when engaging with a new piece of evidence, such as looking up and seeing that it is raining, the properly Bayesian agent will update all of their credences to reflect Bayes' rule: "the prior probability of a belief equals its posterior probability" (Orlandi, 2018, p. 2374). We can then consider a second case where an agent always finds that wet grass is caused by rain without considering the sprinkler system. In contrast to the first case above, the weight for the rain as the cause will steadily increase over time. If the association for the sprinkler system decreases as well, then the system would be acting as a proper Bayesian agent. If the association for the sprinkler remains the same, and even if the association increases for the rain as the cause, then the system would not be acting as a proper Bayes agent.

Whether or not ITPRA can or should be fully interpreted in Bayesian PP terms is an open question that is outside the scope of my current analysis, although I can safely claim that the two are not necessarily opposed theories. In addition to ITPRA, there have been attempts at capturing prediction in music that offer an explicitly PP-based approach to explaining key features of music cognition. Rebecca Schaeffer (2014), for instance, has provided additional information regarding predictive processing as an essential part of musical processing. Schaeffer's main argument concerns the nature of individual responsiveness to music, which she argues comes about when a person applies a mental model, defined as "an internal representation of a percept or action, as it is built up through experience and statistical learning" (2014, p. 161), to a situation. The application of these mental models can also be seen as akin to the application of prior knowledge to predict the incoming flows of sensory information.

In a commentary on Schaefer (2014), John Michael and Thomas Wolf (2014) suggest an expansion of her model to cases of joint action. In addition to pointing out a role for active inference and recursive higher-order modeling to monitor various elements of group performances, they note that PP can be applied to a specific kind of coordination problem regarding how to end a song during an improvised performance. The specific issue here is how several musicians decide when and how to end a song without agreeing on it in advance. In their words, "In the context of music, two jazz performers might both prefer, having just finished what appeared to be the final verse of a jazz standard, to start up again for a surprising additional repetition of the chorus – *but only if the other does so as well*" (2014, p. 180). Dealing with this issue becomes more tractable by a combination of stock endings that are common across a performance community and relying on various embodied and gestural cues to help navigate the end of a performance.

An important upshot from this final point is the fact the PP is not limited to explaining an individual person engaging with the world in isolation. PP provides the basis for describing long-term and short-term, slow and rapid, individual and collective processes of performance over time. In all of these cases, an essential role for expectations and predictions remains.

While this sketch has begun to provide a map for building a broader framework to capture PP in bebop improvisation, it so far leaves untouched the issue of how novelty would come about under a system that prioritizes error minimization. For instance, an additional reading of the jazz coordination problem would be that the best solution for dealing with a complex ending is not just tend towards stock endings, but always use the same ending pattern to ensure a successful completion of the song. This sort of solution would work but would be far from a novel choice. In more extreme terms, and moving back deeper than the application of PP

in musical contexts that already presupposes some potential positive value for eliciting and limiting surprise, a core issue concerns why we even feel positive affect and emotions in relation to novelty in the first place. As such, we shall now turn to facing the issue of novelty head-on, then turn to PP frameworks for jazz improvisation.

PP and Novelty

Novelty and surprise are not coextensive concepts, yet both are important features of improvisation. As such, if PP is to help our understanding of how improvisation functions, it will need to tell part of larger story about novelty. On this particular point, however, a distinct concern becomes clear for PP.

In three steps, this distinct concern runs as follows. First, according to proponents of PP, an organism's main aim is to minimize prediction error, which requires decreasing surprise/surprisal as much as possible. Second, novelty is something that by definition increases surprisal, since, whatever exactly it may be, novelty at least requires being outside the standard statistical regularities faced by an organism. Third, if one and two hold, an organism will do its best to stay away from novelty as much as possible. Unfortunately for supporters of PP, we have clear counterevidence to this final step insofar as people actively seek out novelty in a variety of contexts. Considering the premium placed on novelty in jazz and other artistic practices, novelty may be a central reason why people enjoy playing and listening to jazz and other forms of music in the first place. As a result, supporters of PP are in a *prima facie* bind concerning how to explain novel behavior in bebop performance.

The worry can be extended further since surprise from a biological standpoint is always negative (Huron, 2006). Surprise and surprisal result from a mismatch between expectations of how the world should be and how the world actually is at the moment. Mismatches, in turn, are

not a good thing for an organism that is attempting to get a grip on and survive in a changing world. A mismatch or series of mismatches means that the organism is not gaining an adequate grip on their environment, which generally results in a variety of negative affective responses and, in some cases, may even result in death (e.g. if I am greeted by a hungry tiger instead of my friend at the door, either the tiger or I may die depending on who responds faster to the novel stimuli staring them in the face). As a result, there are at least biologically good reasons for not wanting to experience surprise.

There are several different responses one could give to the general issue of novelty and surprise in biological contexts. One would be to clarify the ambiguity between the conscious, affective feeling of surprise and the subpersonal, statistical notion of surprisal. In itself, the conscious feeling of surprise occurs in both harmful and safe situations. I feel surprised if I arrive at a surprise party in my honor; I likewise would feel surprised if I find a tiger by the front door. The feel of the emotion may not vary between these cases, so appealing to it alone doesn't seem like the best course of action. Of course, there are important phenomenological differences between other parts of these cases, such as joy vs. fear coming along with the respective surprises. There may also need to be additional conceptual work done around the link between such contrasting phenomenological feelings. In either case, it remains the fact that a failure to match expectations alone, even with the attendant phenomenological feeling of surprise, is not automatically a signal that there is biological *threat* facing an organism.

Another response on behalf of PP would be to note that there is a difference between experiencing surprise in the wild and experiencing surprise in controlled situations. Acting out situations within controlled environments, moreover, such as in cases of storytelling, rehearsing, or pretend play, allows us to better prepare for all types of surprise in future situations that could

be stressful and highly dangerous. Fighting mock battles to prepare for the actual experience of war comes to mind for those cases. Other times, experiencing surprise in controlled situations could allow for the thrill of experiencing a threat without it really being a threat, such as going to a haunted house or watching a horror movie. In any of these cases, the feeling of surprise is not a biological threat to the organism, and thus the general fact that mismatch is not good in biological contexts doesn't automatically translate over to it being a threat in bio-social contexts.

In addition to this general problem of novelty and surprise from a biological perspective, however, the concern becomes worse in the case of predictive processing because the PP framework is fundamentally driven by an information-theoretic goal of decreasing surprisal. Adopting either of the proposed solutions above will therefore ultimately prove fruitless for a defender of PP because surprisal would increase across all of those cases, and regardless of whether or not we are considering "good" or "bad" surprise along with it. Therefore, if the main goal of FEM is to decrease surprisal, we are thrown right back into the heart of the quagmire. In response to this new problem, one may suggest that, among the possible environments one could choose to decrease surprisal as much as possible, sitting in a dark, quiet room would be the best option. In a dark room, it should be easy for an organism to predict the future sorts of states they will face and decrease free energy as a result.

If all it took to solve the novelty dilemma was granting an optimal strategy of preferring quiet, dimly lit places over loud, chaotic ones, this solution wouldn't be much of a concern for defenders of PP. A general preference for those sorts of environments does not preclude the ability to leave them from time-to-time, nor does it foreclose the possibility of cultivating an appetite for novelty-seeking or the ability to successfully engage with more vibrant scenarios. However, insofar as minimizing error is the driving principle of an organism, the critic of PP

could further push that dark rooms are not just what an organism prefers but the kind of environment it will actively seek out. We would furthermore not just seek dark rooms out on occasion, but prefer to stay there indefinitely, even until death, without venturing out into the more chaotic world outside. This form of concern has been dubbed the "dark room worry."

There is a simple response to this version of the dark room worry, namely that someone in a constantly dark, quiet, unchanging room would actually be in an extremely surprising situation based on what they have come to expect about how the world works and their attendant biological needs. Constantly being in a dark room is to put oneself in an environment that most organisms would not predict to be normal for their everyday engagements with the world. There are also times when internal homeostatic processes would drive an organism to seek food if they are hungry or drink when they are thirsty. Therefore, while we may not be repelled from enjoying or spending time in dark rooms, there is no reason to assume that minimizing surprisal would be best realized by living in darkness unto death. If correct, this response means that organisms will instead seek out more active environments for various reasons, especially since organisms like us have evolved to expect more complex environments (Friston, Daunizeau, Kilner, & Kiebel, 2010) and expect to be living in environments that are in constant flux (Hohwy, Paton, & Palmer, 2015). Finally, as noted above, the drive to directly address surprisal from states like hunger or thirst can only be fixed by actions (or death), instead of changing the expectant hypotheses about the situation.

There are at least two additional ways to conceive of the dark room worry, however. Following Clark (2018), the first variation canvased above can be called the "Death Trap" version. It's the strongest variation among the dark rooms insofar as it pushes the issue to the point of life or death. The two additional worries are respectively "the Boredom Trap" and the

"Merely Modest Exploration Trap," with the latter being a stronger version of the former. The first, Boredom Trap, is concerned with the idea that an organism will be driven to live in the most predictable environments possible. While not driven to do so to the point of death, they nevertheless prefer to live without any striving for novelty. The second, Merely Modest Exploration Trap, is concerned with the possibility that the fundamental error minimization mechanism of PP can be stuck in such a way that we would run into trouble achieving large-scale positive ends, such as sustained personal growth, true novelty seeking, and general human flourishing.

While these two versions of the dark room are a potential problem across the board, they take an additional level of concern for jazz performance. The evolution of jazz music, as well as most music in general, is driven in a manner largely antithetical to the anti-novelty drive of dark room cases. Considering the extremes that some musicians push themselves for novel innovation, it's clear that they're not settling for some form of Merely Modest exploration either. But how do we account for the fact that these drives for novelty exist?

To start with the weakest dark room variation, and again appealing to Clark, we find that his response to the Boredom Trap is two-fold (2018, p. 527-8). First, turning to the role of affect and neurobiological evidence, Clark notes that there is an important balance of desiring novelty and surprise at the neural level through a complex flow of cortico-subcortical loops. These loops are particularly relevant for reward and value ascriptions. As a result, while surprise may have generally negative biological connotations, those are mitigated by the reward system that positively values novelty. Second, at the species-level, Clark notes that there are benefits for a population of organisms when a subset of individuals deal with known resources in standard ways, while another subset explores new resources and new ways of dealing with old resources.

Although too much novelty-seeking would be bad for a population, not enough novelty would be detrimental to our long-term survival and evolutionary success as well.

While perhaps successful in the case of the Boredom Trap, these responses do not yet guarantee that at least some of the novelty-seeking that a person decides to uptake will be conducive to human flourishing. Hence, we still need to address the Merely Modest Exploration Trap. The core of this trap comes from what Clark has dubbed the threat of "information-theoretic subversion." On the PP account, an organism constantly attempts to minimize prediction error and maximize predictive success by expanding their capabilities to grapple with novelty. However, they may do so in ways that don't track any sort of positive overall flourishing, such as when "an agent could be 'hijacked' by a darkened room containing only a computer monitor and an endless supply of simple (but not too simple) puzzles, each of which allows a steady trajectory of improvement until the next one pops up on the screen" (Clark 2018, p. 530). According to Clark, there is no principled way to develop a response to this concern with PP that is immune to at least some kind of subversion.

The final point in the previous paragraph can be read in its strongest terms as implying that proponents of PP must give up when it comes to novelty, especially if information-theoretic subversion is always possible. Clark suggests, instead, that the reason we rarely fall for this trap is because the humans find themselves in niches that are self-engineered and culturally-laden through and through. A precursor of this reply was introduced above with the fact that humans may use a controlled environment to experience certain kinds of surprise without the biological threats they would pose in other contexts. This point is also the continuation on the idea that humans have built and live in "designer environments."

Such environments, full as they are with various traditions, practices, and material objects, furthermore and continually open up spaces that move us past sedimented forms of cultural practices. They will do so in a way where those who are already members of a community will help scaffold the learning of infants and novices alike, until the less experienced agents become more competent members of the community. Sometimes agents desire to keep things in the community static, while others push for much more novelty. In either case, we have strong cultural and environmental reasons that keep us from eternally returning to information-theoretic subversion.

Another alternative for addressing the Merely Modest Exploration Trap has been proposed by Julian Kiverstein, Mark Miller, and Erik Rietveld (2017). In this case, the key idea for addressing information-theoretic subversion would be to shift the focus from how an organism modulates overall prediction error to how the rate of flow for free energy changes vis-à-vis the organism's predictions regarding how fast it should change. Kiverstein et al. (2017) furthermore suggest that if the prediction error is being reduced at a faster rate than expected, an organism will have a positive emotional response such as joy. In contrast, if error is being reduced at a slower rate than expected, the organism will feel a negative emotional response, such as frustration. Novelty is an essential part of this story since, although novelty opens up the possibility for negative emotions and affects, it is only through novel situations that one is able to experience a greater than expected decrease in the rate of error flow.

We can see another important component here in the role that repetition plays for the development and maintenance of our engagements with the world, which highlights the need for an organism to not go too far into pure novelty seeking at all times. This point is made especially clear in the case of music. Specifically, in the case of bebop, even as the content of the solos may

vary across performances, there is at least some repetition in the head and outro of the song. There is also a different kind of repetition in the solo section, since the underlying chord changes hold across different soloists. As a result of these repetitions, performer and audience members alike are able to distinguish different songs and appreciate the performance that unfolds. Instead of locking ourselves into a dark room, we instead fruitfully combine repetition and novelty, just as the practice of improvisation in general combines novelty with preexistent components.

The critic may retort here that repetition, while useful across cases, remains at its core antithetical to novelty. This worry is particularly acute if enough repetition morphs a behavior into habit and/or threatens to make it dull and uninteresting. To the first point, I return again to my original definition of improvisation as a mix of plans and spontaneity. Repetition and novelty ultimately go together here in the long run. In response to the second point, there has been a wealth of empirical evidence gathered in favor of what Robert Zajonc originally dubbed the "mere exposure effect," and which Huron (2006, Chap 8 & 13) develops further for music as a type of predictive effect. As the original name implies, the mere exposure effect rests on the fact that simply being exposed to something repeatedly will increase preference for it, instead of automatically making it dull and uninteresting.

Musicians are adept at utilizing a prediction effect to relate to listeners as well. Huron (2006, p. 230) highlights four different basic paths to this goal: schematic predictability (conforming to existent listening schemas), dynamic predictability (work is structured to evoke work-specific expectations), veridical familiarity (listen to the same piece many, many, many times), and conscious predictability (a knowledgeable listener is able to consciously predict what will happen as the music progresses). Jazz musicians, as with all other musicians, may draw on one or more of these four basic tools.

Many cases of jazz performance will break with the specifics of these different methods, however. For instance, while the underlying form of a standard remains the same across contexts, it can vary greatly between different performances, making strict veridical familiarity difficult to come by across performances. There are likewise relatively few work-specific properties in jazz standards, which has an impact on dynamic predictability. Finally, based on the inherent role of surprise, being too consciously predictable or holding too closely to schematic predictability will be detrimental as well, since these options lack the novelty both under consideration here and prized by jazz performers and fans alike.

There are two additional points raised by Huron that are relevant to considering novelty in this current context. First, the preceding discussion does not take into consideration whether the resulting music would be "good" or not. There are additional considerations that distinguish good music from not-so-good music, which would explain why the four basic methods for developing predictability are not exhaustive for making good jazz music, as well as the reason why we cannot simply appeal to them in order to establish high-quality jazz performances. Second, and following from this final point, the predictive effect is not the only relevant psychological mechanism or process at play. These other psychological components will thus limit or enhance the overall positive emotional impact of repetition.

Returning to the dark room idea plus the possibility to appeal to further psychological mechanisms, we are here faced with a final consideration about error minimization in PP. Put directly, does accepting either Clark's designer environment or Kiverstein et al.'s variable rate of surprisal require rethinking the fundamental nature of PP and the information-theoretic demand to reduce prediction error? While some people have suggested that it does, Clark has argued that it does not. Instead, in addition to cultural and neurobiological reasons, humans avoid the dark

room traps for homeostatic reasons – i.e. hunger and thirst drive us to act – as well as the fact that it is only through active inference in the world that we will be able to decrease surprisal in a world that is far from always being dark and calm.

While both of these proposed solutions – designer environments and rates of rates of change - can stand and fall on their own merits, I suggest that the Cultural Affordances Framework offers one way for bringing elements of them together. I will thus now turn to introducing the CAF to see how it handles this connection as well as a different proposed solution to the dark room worries.

A Cultural Affordances Framework – Bridging Ecological and PP Approaches

When first introducing predictive processing, I briefly considered one framework – the Active Inference Framework (AIF) – that has been offered for bringing PP into contact with other research areas. While that framework is within the bounds of 4E cognitive science, it stops short of the more radical camps as developed in enactivism. In this section, I will consider the Cultural Affordances Framework (CAF), which is closer in spirit to various tenants of the enactivist paradigm. This framework also offers a unique take on the role that other minds play in shaping the various predictions that an agent utilizes during cognition.

One of the major theoretical benefits of the CAF is that it places a priority on the fact that humans are enculturated within "regimes of shared attention" and "regimes of expectation." A driving part of the CAF story is that cognition cannot be limited to what the individual expects to happen in the world. Instead, expectation, both in general and specifically in regards to predictive processing, has an intrinsically social aspect that is cultivated through our lifelong engagement with others. A main upshot of the CAF therefore encompasses broader normative and inherently social factors in the development of priors and expectations. These social factors

are also a foundational part of cognition. This social link is operant so much so that Veissière et al. (Forthcoming) refer to the underlying intersubjective processes in cognition as moving beyond an account of Theory of Mind to the idea the mind is fundamentally grounded in "Thinking through Other Minds" (TTOM).

TTOM and its attendant social components – referred to as "culturally patterned practices" in more technical terms – have an impact on all elements of PP, including predictions, prediction error, precision weighting, and the modulation of attention (Ramstead et al., 2016, p. 12). These influences start in early childhood, when caregivers make an impact on infant emotional responses (Hanford et al., 2018) and can be witnessed in the ability for adults to utilize expectations to modulate behavior when moving across different social contexts.

TTOM is an explicitly metacognitive theory insofar as it requires an individual to track the cognitive states of oneself and other agents, especially other people, in order to (at least implicitly) model them and respond to these expectations of others as needed. Considering the general cognitive robustness required to utilize metacognitive resources, one could here ask how does the framework begin for infants who have yet to develop such complex cognitive capabilities?

In reply, the claim from proponents of TTOM is to focus on the developmental process underlying human cognition. Infants are born already primed to take part in social activities. Over time, it is the actual social activity itself that gives rise to each individual set of expectations of the world. In other words, while all infants are primed to turn towards caregivers to make sense of what is good and bad in the world, how they develop those attendant expectations will be idiosyncratic and based on context relevant learning processes. The story of

TTOM therefore has the tools to both explain how the process of learning gets of the ground and explain how it develops over the course of an individual's lifetime.

Among its different benefits, the CAF is a particularly appealing theory when used to understand either the process of becoming a musician or the process of how a particular musical idiom develops. Learning to become a musician occurs within regimes of shared attention, after all. These regimes can be formal or informal institutions, from college classrooms to jazz club jams. Such regimes also teach us what it is we should (or should not) be focusing on. Put in other words, the development of genres can be traced to developments and differences among regimes of expectations.

Is it possible for an agent to stumble onto a variety of musical affordances when engaging with aspects of their environment outside of being explicitly enculturated thorough regimes of shared attention? That may be so; however, both passing on skills through shared attention and developing a distinctly normative understanding of how these skills relate with success and quality cannot be explained in the lucky and socially isolated agent. Cultural learning likewise helps make the process of cultural transmission faster, more efficient, and more lasting.

In developing the CAF, Maxwell Ramstead, Samuel Veissière, and Laurence Kirmayer (2016) make a key additional distinction among types of affordances. Instead of only appealing to cultural affordances, they suggest that we can trace differences between natural affordances and conventional affordances. Drawing on the work of Paul Grice, these distinct types trade on natural and non-natural information, respectively.[5] Natural information is a non-semantic kind of information that relies on reliable correlations, while non-natural information involves norms as

[5] This distinction nicely mirrors the distinction between lawlike and general ecological information as introduced by Bruineberg, Chemero, & Rietveld (2018). An important difference between these two ways of carving up affordance space, in addition to the use of Grice here, is that the CAF focuses on the case of learning affordances, while the general/lawlike distinction is more focused on the general type of information that can be learned by an organism.

well as explicit and implicit background cultural knowledge and information. It is possible to engage with natural information without taking part in metacognitive monitoring of others, although their presence may have an impact on what affordances are salient in the field of affordances at a given time. In contrast, it is essential for an agent to master the expectations behind cultural scripts, norms, codes, and knowledge to leverage conventional affordances.

Music is a clear paradigm phenomenon that requires both conventional and natural affordances. The raw material required to make a jazz solo is replete with a variety of natural affordances. A particular tuning on a guitar accompanied with a distinct plucking pattern will create tones and follow basic laws of sound. While these sounds are created by and for humans, even if there weren't cultural critters like us around to make use of those sounds after they were created, they would presumably still develop in the same manner. At the same time, how soloists choose to use or ignore different natural affordances are laden with socio-cultural meanings that are better understood in the realm of conventional affordances. Some conventional affordances can be seen in practices such as quoting different licks from other musicians. Others include making reference to previous solos from either the same show or another time, which is often a source of great delight and importance for performers and audience alike.

While these above examples provide information for describing important parts of music cognition, they do not yet provide the tools to answer the primary question of this chapter regarding novelty. How does the addition of regimes of shared attention, regimes of expectations, conventional affordances, TTOM, and the like help us explain novelty in jazz performance in a way that was not possible before? Part of the answer can be gleaned from how Samuel Veissière, Axel Constant. Maxwell Ramstead, Karl Friston, and Laurence Kirmayer (Forthcoming) offer an answer to the dark room worries. In their words:

> agents will act to optimise the *epistemic value or affordance* of an action before acting on its pragmatic value, which is essentially its expected utility. For example, if one enters a dimly lit kitchen to grab a midnight snack from the pantry, one is more likely to turn the light switch on before heading to the pantry. Turning the light on allows one to get an optimal grip and disambiguate the situation, before one acts on the pragmatic value (i.e., the utility) offered by snack foods. (p. 39, emphasis in original).

In this example, there is an essential link between reduction of uncertainty and exploration. Such a link is also inherent to the role of FEM, since the drive to minimize uncertainty pairs well with a need to explore the environment. Such exploration may, they note, lead to a local increase in free energy with an end result of a greater decrease of free energy in the long-term.

However, in the case of jazz performance, it's not clear that the privileging of epistemic value is always acted on before pragmatic value or expected utility, at least if epistemic value is taken as fully disambiguating the value of an action before acting on the expectant pragmatic values. Furthermore, while this reply to the dark room worry provides reasons for accepting how CAF can address novelty, it seems to fall short in a way. The primary failure here, as acknowledged by Veissière et al. (Forthcoming, p. 63), comes from the fact that the CAF takes a conservative approach to culture. This conservatism can be understood in the sense that their account of enculturation, like many others, presumes a "stable social context." Such an assumption is a common point of agreement across many accounts of culture and enculturation, as the authors note (p. 63), yet social change generally does not take place in such stable environments. The case of when and how bebop jazz developed shows one such limitation of this account, since not only can it be placed at an interesting juncture within and among different

cultural forces (e.g. the combination of black and white musicians performing across a segregated US, during and shortly after WWII, along with the various economic pressures from gigging and the aesthetic pressures to develop new styles of playing that kept the music fresh and alive for listeners and performers alike), but it also offers a wealth of historical data to trace out how audience and performer practices shifted within these various cultural changes. In addition, exploring these historical roots remain important today, even as parts of the contemporary performance practice around bebop has trended towards a conservative view of its canon.

It is on this point that putting CAF and bebop improvisation in conversation shows a future for benefits flowing from the latter to the former. Following the call for a focus on cultural mobility and hybridity, in addition to providing a case study for explaining some of the technical terms of the CAF, I suggest that by tracing out the particular developments of bebop, socially and musically, can help establish a clearer story of how unstable social situations can develop conflicting regimes of attention and expectations, as well as how people have navigated through those conflicts.[6] While providing this story is outside the scope of my current project, it remains the case that CAF has provided some of the basic tools that can be further augmented and developed moving forward in understanding jazz improvisation as well as the development of other musical genres and subgenres.

[6] There is an additional upshot here concerning DPM model from Chapter 2 and the CAF. The latter has a strong story to tell at the general, social level – with the caveats noted in the previous paragraph notwithstanding – of jazz performance, along with the role of predictive processing in the brain. It also has a budding story to tell about the role that expectations of others play in Thinking through Other Minds (TTOM). What it's missing at the group level, however, is an explanation of how the intentional activity of individuals in groups after they have developed within regimes of shared attention and regimes of expectation. The dynamical flow of individual and shared intentions of the DPM model thus provides an additional explanatory tool to flesh out this intermediate level between the explanation that the CAF provides at the neural level with PP and at the social level with an account of culture.

Chapter Conclusion

In this chapter, I considered the role of expectation in music, novelty jazz improvisation, the problem of novelty in predictive processing, and how several predictive processing frameworks can offer a general account that brings these various components, and more, together under one theoretical banner. In addition to considering and responding to general worries about how PP and free energy minimization can be harmonized with novelty, I proposed the CAF as a possible framework that could help us better understand jazz. I also suggested that the CAF would benefit from taking bebop jazz as a case study that pushes the framework outside of its conservative tendencies towards culture and instead moves it towards taking cultural change, hybridity, and cultural differences as the norm to be better understood and explained.

Conclusion

In this dissertation, I have considered how tools from embodied cognitive science may be utilized to better understand the cognitive processes underlying bebop performance. The edifice that I have constructed, spanning across shared intentions, ecological psychology, and predictive processing, requires sources that are both empirical and theoretical in origin. The overall structure of this project - from the conceptual analysis of what constitutes improvisation, during which I suggested both a general definition as well as one specific to the case of jazz, through the framework of cultural affordances and cultural learning – has, in part, developed as its own sort of improvisational dance between shared intentions, ecological psychology, and predictive processing. In working through this full story, as well as the ensuing considerations about these three research areas, I believe that three general themes returned in different guises throughout the dissertation. In addition, I will briefly consider how all of these points can speak to existent models of jazz improvisation.

First, a consistent theme across the ideas canvassed above was the continual difficulty that many accounts faced concerning uncertain stimuli that arose from the flexibility of the bebop jazz performance environment. Some of these problems developed because the theory in question heretofore assumed that the environment is almost entirely unambiguous, at least in the relevant manner that what is tracked by the various perceptual systems, aimed as they are at "getting a grip" on the world, can be understood against the backdrop of an environment that is relatively stable, all things considered, and in which people have developed to fit those environments through changes at both ontogenetic and phylogenetic timescale. While part of jazz performance is about fitting performance expectations, it cannot be limited in that same way. Other theories ran into the problem of uncertainty by considering the environment as too

random, in which case they did not grant enough role to the fact that improvisation relies on existent forms of action, schemes, and memories as well as developing a sense of novelty.

The theme around uncertainty and ambiguity, in turn, served two purposes in different areas. For one, I suggested it as a motivation for developing new accounts and pushing old theories into new directions. Second, it occasionally served to directly critique some positions on offer in the literature. In either case, it's clear that we still have a long way to go to account for this sort of uncertainty underlying improvisation.

Second, what novelty is, and how it can be achieved, was an important consideration throughout this project. The exact answer to this question varied depending on whether or not we were considering experts or novices, and whether or not we were considering the case of individual or group processes. Like uncertainty, it also at times was the target I aimed to explain while, at others, it was a concern that impacted my analysis of different accounts.

Third, another issue concerned the ways in which improvisation is (or is not) supported by lawlike explanations. From Elisabeth Pacherie's concern that laws fail in the case of action explanations to the ecological psychologist invocation of a lawlike account of perception, I considered positions that were critical of lawlike explanations as well as those that were in favor of them. Similar to the cases of uncertainty and novelty, jazz performance is useful for these debates insofar as it offers a rich, ecological context that may ultimately call for splitting the difference between these two extremes. Certain parts of musical performance will follow law-like components while others will not. This position likewise speaks well of continued development for large, interdisciplinary and transdisciplinary frameworks to fully understand what is happening in the case of the cognition occurring when musicians improvise a performance, as well as in the case of improvisation in general.

I will now turn to offer a brief sketch of how the three areas of research considered above – shared intentions, ecological psychology, and predictive processing – can say something about the cognitive models for musicians-in-action. To do so, I will first briefly introduce the model of improvisation provided by Jeff Pressing (1988), which is one of the most influential to date, along with an ecological description of jazz improvisation from Stefan Carris Love (2017). I take my work to help extend both the model and description, as well as open up possible new avenues of research moving forward.

According to Pressing (1988), a model of improvisation must give a story concerning (1) the origins of novel behavior in improvisation, (2) how people develop improvisational skills, and (3) an account of how people actually improvise. An important additional distinction across these three *explananda* is made between improvisations that involve a referent and those that are referent-free. Pressing endorses a broad definition concerning what counts as referent, including any sort of underlying schema or guide that a musician can utilize while crafting their improvisations. In the case of bebop, this would include the song being played, the chord changes specified for the solo, as well as various additional knowledge around that particular piece or about the other musicians in an ensemble performance. It could also include the use of a lead sheet (if one is available and utilized). As a result, the referent can include material, environmental, and internal tools to ground and help facilitate the act of improvisation.

Pressing breaks any improvisation into a series of "event clusters," each of which has several aspects. Aspects can further be broken into three parts: objects, features, and processes. An object is any relevant entity that is the target of cognition or perception, such as a particular chord or a fingering pattern. Features bridge together the parameters that are shared across different objects. Processes concern the changes of both objects and features over time. When

moving between different event clusters, a musician often decides to either build additional associations or interrupt the previous continuity of event clusters.

During associative generation, the musician maintains many of the existent components from previous object, feature, and property arrays, and decides on where to build from them. In interruption generation, the musician decides to break with one, many, or all previous parameters and replace them with new ones. In an ensemble performance, it is possible that other musicians may also support different associative generations or demand interruptive generations. This story also entails some considerations about how novelty is developed during improvisation. As noted by Pressing, this can include developing movement control parameters for newly discovered objects, processes, or features, or enacting novel combinations between the components.

Love (2017) has developed an ecological description of jazz improvisation that took Pressing's model as a main starting point. Combining it with elements of ecological psychology, however, he offers several important changes. For instance, Love's description puts pressure on the computational nature of Pressing's account, which relied on making clear distinctions among different event clusters and the unfolding relationship between them. Moreover, Love suggests three main areas that are changed between Pressing's model and his account.

First, the nature of the referent is reconsidered to not only make improvisation easier but also more difficult. The difficulty comes from the fact that the referent both opens and forecloses available choices. This foreclosure process, in turn, makes playing a jazz improvisation like "dancing through an obstacle course" (Love, 2017, p. 34). Such obstacles are furthermore not reducible to music-theoretical terms, although music-theoretical components will be included as an important part of the musical environment that is being navigated. Second, instead of pushing the majority of learning into the realm of developing a store of memorized objects or rules, the

ecological account focuses our attention on changes to perceptual processes and the nature of the musical environment in which one is navigating their solo. Third, to account for the speed and timing of improvisation, Love suggests that we refocus from a question of "when" to a question of "where" (2017, p. 40). The musical environment and a soloist's anticipations about future events cannot be limited to a "snapshot" of time but includes the overall flow of perception that is always taking place in a particular performance environment.

My work lines up well with all three of Love's alterations to Pressing's account. Love's use of navigation as a guiding metaphor dovetails closely with my general and specific definitions of improvisation and jazz improvisation, for instance. I also introduce two additional layers that are not considered in his ecological description. The first concerns a clearer role for the brain, both in regards to resonance and predictive processing. The second considers the dynamics of ensemble performance through the lens of shared intentions. One additional point we diverge on concerns the nature and role of "errors" in jazz performance. However, that disagreement is not directly related to the issue of modeling at hand and, instead, concerns other considerations about what makes a good or successful jazz improvisation.

Furthermore, the final consideration about frameworks from the previous chapter also points to a further way in which these two accounts, and perhaps other cognitive models of improvisation, may be synthesized into a more general account of jazz improvisation, as well as improvisation in general, in both solo and collective contexts. I suggest that one way for doing so would look to the Active Inference Framework (AIF) and the Cultural Affordance Framework (CAF), both in terms of an important guideline for how this unification process can occur and for being taken into developing the AIF, CAF, as well as the Skilled Intentionality Framework (SIF) in the future.

Across all of these cases, I suggest that my focus on bebop improvisation ultimately serves a dual purpose. On one hand, it helps flesh out certain principles in action that brings several general theories into heretofore underexplored areas. For instance, in the AIF, the nature of interoception, proprioception, and exteroception is distinct in the case of jazz improvisation from other sorts of improvisational activities. How these differences will manifest is an open empirical question. In the case of the CAF, for example, jazz requires and simultaneously offers a space for understanding culture itself across variable changes, instead of simply offering another case of presupposing stability and conservatism in cultural development. In both of these cases, jazz is a place where the given accounts can be further developed in the future.

On the other hand, bebop improvisation as the target of analysis can be understood through appeal to work in the CAF, SIF, and AIF, as well as related areas. A version of this path is seen in the way Love's ecological description appeals to ecological psychology to better understand jazz improvisation, instead of seeing how a focus on jazz improvisation can change our understanding of ecological psychology.

While jazz performances are crafted to open up improvisational skills and opportunities, the act of improvising itself is not only found in such highly specialized situations. People take part in improvised activities during their everyday engagements with the world as well. From navigating traffic jams or public transportation to chatting with strangers and friends alike, most daily activities require on-the-fly adaptation to changes in the environment. Although my dissertation has primarily been focused on one particular form of a skilled improvisational practice, I believe that there are broad insights that can be applied to other areas as well.

Since bebop jazz offers a case that is more constrained than free jazz yet more free than swing era jazz, its usefulness as a case study may also be a greater chance for general cognitive

overlaps between it and other cases of improvisation, While this possibility is ultimately an empirical question, if it were the case, then the suite of capabilities highlighted throughout the dissertation may be useful in guiding future empirical work into areas that are pedagogical and practical across a wide domain of actions.

Moreover, while jazz improvisation may or may not be the best metaphor for understanding the brain and cognition, I hope to have at least shown that it, as well as other genres of music, not only deserves a much larger place in the topics under consideration in cognitive science – a space in which, independently from this project, I am happy to see it starting to appear with greater speed in recent years (e.g. Torrance & Schumann, 2019) – but that its study cannot be reduced to rehashing the same things we have already learned from the domain of language or other sorts of skilled activities.

I started this dissertation from the sense of wonder that comes from experiencing masters of improvisation in the midst of performance. The wonder remains, as I believe at least some part of it always will, but I hope to have contributed something to making a bit more sense of how such a wonderous activity can be realized in practice. Indeed, the ultimate cognitive story that we will tell about jazz improvisation, towards which this dissertation may take only several small steps, will not crush that wonder. It will instead, I hope, allow more people to embrace it and find a deeper sense of joy that comes from learning some of the secrets behind how improvisation is possible, and, perhaps, using those skills in other parts of their lives as well.

Works Cited

Alperson, P. (2007). A topography of improvisation. *Journal of Aesthetics and Art Criticism, 68,* 273-279.

Alperson, P. (2017). Musical improvisation and the philosophy of music. In G.E. Lewis & B. Piekut. (Eds.), *The Oxford handbook of critical studies. Vol. 1* (pp. 419-438). Oxford U. Press.

Ashley, R.D. (2008). Musical improvisation. In S, Hallam, I. Cross, and M. Thaut (Eds.), *The Oxford handbook of music psychology* (pp. 413-420). Oxford U. Press.

Barrett, L., & Bar, M. (2009). See it with feeling: affective predictions during object perception. *Philosophical Transactions of the Royal Society of London B Biological Sciences, 364,* 1325-1334.

Beaty, R.E. (2015). The neuroscience of musical improvisation. *Neuroscience & Biobehavioral Reviews, 51,* 108-117.

Becchio, C., Manera, V., Sartori, L., Cavallo, A., & Castiello U. (2012). Grasping intentions: from thought experiments to empirical evidence. *Frontiers in Human Neuroscience, 6,* 1-6.

Bengtsson, S.L. Csíkszentmihályi, M., & Ullén, F. (2007). Cortical regions involved in the generation of musical structures during improvisation in pianists. *Journal of Cognitive Neuroscience. 19,* 830-42.

Berkowitz, A.L., & Ansari, D. (2008). Generation of novel motor sequences: the neural correlates of musical improvisation. *Neuroimage, 41,* 535-43.

Berliner, P. (1994). *Thinking in jazz: The infinite art of improvisation*. U Chicago Press.

Borgo, D. (2005). *Sync or swarm: Improvising music in a complex age*. New York: Continuum.

Bratman, M.E. (1987). *Intentions, plans, and practical reason.* Cambridge, MA: Harvard U. Press.

Bratman, M.E. (1993). Shared intention. *Ethics, 104,* 97-113.

Bratman, M.E. (1999). *Faces of intention: Selected essays on intention and agency.* Cambridge, UK: Cambridge U. Press.

Bratman, M.E. (2014). *Shared agency: A planning theory of acting together.* Oxford, UK: Oxford U. Press.

Brown, L.B. (1996). Musical works, improvisation, and the principle of continuity. *The Journal of Aesthetics and Art Criticism*, *54*, 353-369.

Brown, L.B. Goldblatt, D., & Gracyk, T. (2018). *Jazz and the philosophy of art*. Routledge.

Brozzo, C. (2017). Motor intentions: How intentions and motor representations come together. *Mind & Language, 32,* 231-256.

Bruineberg, J., Chemero, A., & Rietveld, E. (2018). General ecological information supports engagement with affordances for 'higher' cognition. *Synthese* Online Version, https://doi.org/10.1007/s11229-018-1716-9

Bruineberg, J., Kiverstein, J., & Rietveld, E. (2018). The anticipating brain is not a scientist: the free-energy principle from an ecological-enactive perspective. *Synthese, 195*, 2417-2444.

Burland, K., & Windsor, W.L. (2014). Moving the gong: exploring the contexts of improvisation and composition. In K. Burland & S. Pitts (Eds.), *Coughing and clapping: Investigating audience experience* (pp. 101-114). Ashgate.

Caiani, S.Z. (2014). Extending the notion of affordance. *Phenomenology and the Cognitive Scienc*es, *13*, 275–293.

Carroll, J.W. (2016). Laws of nature. *The Stanford Encyclopedia of Philosophy*, E.N. Zalta (Ed.).

Carter, A. (1990). On individualism, collectivism, and interrelationism. *Heythrop Journal, 31,* 23-38.

Chemero, A. (2003). An outline of a theory of affordances. *Ecological Psychology, 15*, 181-195.

Chemero, A. (2009). *Radical embodied cognitive science.* Cambridge, MA: MIT Press.

Chemero, A., & Turvey, M.T. (2007). Gibsonian affordances for roboticists. *Adaptive Behavior*, 473-480.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1980). *Knowledge of language: Its nature, origin, and use.* New York: Prager.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36,* 1-73.

Clark, A. (2015). Radical predictive processing. *The Southern Journal of Philosophy, 53,* 3-27.

Clark, A. (2016). *Surfing uncertainty: Prediction, action, and the embodied mind*. Oxford, UK: Oxford U. Press.

Clark, A. (2018). A nice surprise? Predictive processing and the active pursuit of novelty. *Phenomenology and the Cognitive Sciences, 17*, 521-534.

Clarke, E.F. (2005). *Ways of listening: An ecological approach to perception of musical meaning*. Oxford, UK: Oxford U. Press.

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

Cross, I. (2001). The evolutionary nature of musical meaning. *Musicae Scientiae, Special Issue*. 179-200.

Davidson, D. (1963). Actions, reasons, and causes. *Journal of Philosophy, 60*, 685-700.

Davidson, G.L., Cooke, A.C., Johnson, C.N., & Quinn, J.L. (2018). The gut microbiome as a driver of individual variation in cognition and functional behaviour. *Philosophical Transactions of the Royal Society B Biological Sciences, 373*, doi: 10.1098/rstb.2017.0286.

De Bruin, L.R. (2015). Theory and practice in idea generation and creativity in jazz improvisation. *Australian Journal of Music Education, 1*, 91-115.

De Bruin, L.R., & Harris, A. (2017). Fostering creative ecologies in australasian secondary schools. *Australian Journal of Teacher Education, 42*, 23-43.

DeNora, T. (2000). *Music in everyday life*. Cambridge, UK: Cambridge U. Press.

De Souza, J. (2017). *Music at hand: Instruments, bodies, and cognition*. Oxford U. Press.

Deutsch, D. (2013). *The psychology of music.* D. Deutsch (Ed.). (3rd ed.). Academic Press.

DeVeaux, S. (1997). *The birth of bebop: A social and musical history*. U. of California Press.

Dibben, N. (2001). What do we hear, when we hear music?: Music perception and musical material. *Musicae Scientiae, V*, 161-194.

Drayson, Z. (2017). Modularity and the predictive mind. In T. Metzinger & W. Weise (Eds.), *Philosophy and Predictive Processing.* Frankfurt am Main: MIND Group.

Elliot, D., & Silverman, M. (2015). *Music matters: A philosophy of music education*. (2nd ed.). New York: Oxford U. Press.

Engel, A.K., Fries, P., & Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top-down processing. *Nature Review of Neuroscience, 2*, 704-716.

Epstein, B. (2018). Social ontology. *The Stanford encyclopedia of philosophy* E.N. Zalta (Ed.).

Fadiga L, Craighero L., & D'Ausilio, A. (2009). Broca's area in language, action, and music. *Annals of the NY Academy of Sciences, 1169*, 448-58.

Fodor, J.A. (1980). Methodological solipsism considered as a research strategy in cognitive science. *Behavioral And Brain Sciences, 3*, 63-109.

Fodor, J.A. (1983). *The modularity of mind.* Cambridge, MA: MIT Press.

Fodor, J.A. (2008). *LOT 2: The language of thought revisited*. Oxford, UK: Oxford U. Press.

Friston, K. (2012). The history of the future of the Bayesian brain. *Neuroimage, 62*, 1230-1233.

Friston, K. (2013). Life as we know it. *Journal of The Royal Society Interface*, *10*, 20130475.

Friston, K.J., Daunizeau, J., Kilner, J., & Kiebel, S.J. (2010). Action and behavior: a free energy formulation. *Biological Cybernetics*. Online Version, DOI 10.1007/s00422-010-0364-z.

Fuchs, T. (2018). *Ecology of the brain: The phenomenology and biology of the embodied mind*. Oxford, UK: Oxford U. Press.

Gallagher, S. (2001). The practice of mind: Theory, simulation or primary interaction? *Journal of Consciousness Studies, 8*, 83-108.

Gallagher, S. (2008). Direct perception in the intersubjective context. *Consciousness and Cognition, 17*, 535-543.

Gallagher, S. (2017). *Enactivist interventions: Rethinking the mind*. Oxford, UK: Oxford U. Press.

Gallagher, S., & Allen, M. (2018). Active Inference, enactivism and the hermeneutics of social cognition. *Synthese, 195*, 2627-2648.

Gallese, V. (2005). "Being like me": Self-other identity, mirror neurons, and empathy. In S. Hurley & N. Chater (Eds.), *Perspectives on imitation: From neuroscience to social science, Vol. 1. Mechanisms of imitation and imitation in animals* (pp. 101-118). Cambridge, MA: MIT Press.

Gaver, W.W. (1993a). What in the world do we hear? An ecological approach to auditory event perception. *Ecological Psychology, 5*, 1-29.

Gaver, W.W. (1993b). How do we hear in the world? Explorations in ecological acoustics. *Ecological Psychology, 5*, 285-313.

Gibson, J.J. (1966). *The senses considered as perceptual systems.* Oxford, UK: Houghton Mifflin.

Gibson, J.J. (1979). *The ecological approach to visual perception.* Houghton Mifflin.

Gilbert, M. (1987). Modelling collective belief. *Synthese, 73*, 185-204.

Gilbert, M. (1989). *On social facts.* New York: Routledge.

Gilbert, M. (1990). Walking together: A paradigmatic social phenomenon. *Midwest Studies in Philosophy, 15*, 1-14.

Gilbert, M. (1994). Remarks on collective belief. In F. Schmitt (Ed.) *Socializing epistemology: The social dimension of knowledge*. (pp. 235-255). Maryland: Rowman & Littlefield.

Gilbert, M. (2004). Collective epistemology. *Episteme, 1*, 95-107.

Gilbert, M. (2014). *Joint commitment: How we make the social world.* Oxford U. Press.

Gioia, T. (1988). *The imperfect art: Reflections on jazz and modern culture*. Oxford U. Press.

Gioia, T. (2012). *The jazz standards: A guide to the repertoire*. Oxford: Oxford U. Press.

Goldman, A.J. (2016). Improvisation as a way of knowing. *Music Theory Online, 22*, 4.

Gould, C.S., & Keaton, K. (2000). The essential role of improvisation in musical performance. *The Journal of Aesthetics and Art Criticism, 58*, 143-148.

Gratier, M. (2008). Grounding in musical interaction: Evidence from jazz performances. *Musicae Scientiae, Special Issue*, 71-110.

Hagberg, G. (2016). Ensemble improvisation, collective intention, and group attention. In G.E. Lewis & B. Piekut (Eds.), *The Oxford handbook of critical improvisation studies, Vol. 1* (pp. 481-499). Oxford U. Press.

Hagberg, G. (2017). The ensemble as plural subject: jazz improvisation, collective intention, and group agency. In E.F. Clarke & M. Doffman (Eds.), *Distributed creativity: collaboration and improvisation in contemporary music* (pp. 300–313). Oxford U.Press.

Hanford, L.C., Schmithorst, V.J., Panigrahy, A., Lee, V., Ridley, J., Bonar, L.,…& Phillips, M.L. (2018). The impact of caregiving on the association between infant emotional behavior and resting state neural network functional topology. *Frontiers in Psychology*, https://doi.org/10.3389/fpsyg.2018.01968.

Hatch, M.J., & Weick, K.E. (1998). Critical resistance to the jazz metaphor. *Organization Science, 9*, 600-604.

Heft, H. (2001). *Ecological psychology in context: James Gibson, Roger Barker, and the legacy of William James's radical empiricism*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Heras-Escribano, M. (2017). Non-factualist dispositionalism. *Philosophia, 45*, 607-629.

Heras-Escribano, M. (2019). *The philosophy of affordances*. Palgrave MacMillan.

Hirose, N. (2011). Affordances, effectivities, and extension of the body. In W. Tschacher & C. Bergomi (Eds.), *The implications of embodiment: cognition and communication.* (pp. 231-252). Imprint.

Høffding, S. (2019). *A phenomenology of musical absorption*. Palgrave Macmillan.

Hohwy, J. (2013). *The predictive mind.* Oxford U. Press.

Hohwy, J., Paton, B., & Palmer, C. (2015). Distrusting the present. *Phenomenology and the Cognitive Sciences, 15*, 315-335.

Hurley, S. (2008). The shared circuits model (SCM): How control, mirroring, and simulation can enable imitation, deliberation, and mindreading. *Behavioral and Brain Sciences, 31*, 1-58.

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

Hutto, D. (2016). Getting into predictive processing's great guessing game: Bootstrap heaven or hell? *Synthese, 195*, 2445-2458.

Hutto, D.D., & Myin, E. (2017). *Evolving enactivism: Basic minds meet content.* Cambridge, MA: MIT Press.

Iacoboni M., Molnar-Szakacs I., Gallese V., Buccino G., Mazziotta J.C., & Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLOS Biology, 3*, e79.

Ingold, T. (2000). *The Perception of the Environment: Essays on livelihood, dwelling and skill.* New York: Routledge.

Iyer, V. (2004). Improvisation, temporality and embodied experience. *Journal of Consciousness Studies, 11*, 159-173.

Johnson-Laird, P.N. (2002). How jazz musicians improvise. *Music Perception, 19*, 415-442.

Johnston, J., Seibel, S., Darnet, L.S.A., Renninger, S., Orger, M., & Lagnado, M.L. (2019). A retinal circuit generating a dynamic predictive code for oriented features. *Neuron, 102*, 1211-1222.

Kamoche, K., e Cunha, M.P., & da Cunha, J.V. (2003). Towards a theory of organizational improvisation: Looking beyond the jazz metaphor. *Journal of Management Studies, 40*, 2023-2051.

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., Novembre, G., & Loehr, J. (2016). Musical ensemble performance: representing self, other, and joint action outcomes. In S.S. Obhi & E.S. Cross (Eds.), *Shared representations*: *Sensorimotor foundations of social life* (pp. 280-310). Cambridge U. Press.

Kello C.T., Bella S.D., Médé B., & Balasubramaniam, R. (2017). Hierarchical temporal structure in music, speech and animal vocalizations: jazz is like a conversation, humpbacks sing like hermit thrushes. *Journal of the Royal Society Interface, 14*, 20170231.

Kersten, L. (2017). Extended music cognition. *Review of Philosophy and Psychology, 8*, 501-517.

Kiverstein, J., & Clark, A. (2009). Introduction: Mind embodied, embedded, enacted: One church or many? *Topoi*, *28*, 1-7.

Kiverstein, J., Miller, M., & Rietveld, E. (2017). The feeling of grip: novelty, error dynamics, and the predictive brain. *Synthese,* Online Version, doi 10.1007/s11229-017-1583-9.

Knoblich, G., Butterfill, S., & Sebanz, N. (2011). Psychological research on joint action: Theory and data. In B. Ross (Ed.), *The psychology of learning and motivation, Vol. 54* (pp. 59-101). Burlington: Academic Press.

Kourtis, D., Knoblich, G., Woźniak, M., & Sebanz, N. (2014). Attention allocation and task representation during joint action planning. *Journal of Cognitive Neuroscience, 26*, 2275–2286.

Krueger, J. (2014). Affordances and the musically extended mind. *Frontiers in Psychology*, *4*, 1-13.

Krueger, J. (2009). Enacting musical experience. *Journal of Consciousness Studies, 16*, 98-123.

Krueger, J. (2011). Doing things with music. *Phenomenology and the Cognitive Sciences, 10*, 1-22.

Krumhansl, C.L. (1990). *Cognitive foundations of musical pitch*. Oxford, UK: Oxford U. Press.

Laland, K., Matthews, B. & Feldman, M.W. (2016). An introduction to niche construction theory. *Evolutionary Ecology, 30*, 191–202.

Laland, K.N, Odling-Smee, J. & Feldman, M.W. (2000). Niche construction, biological evolution, and cultural change. *Behavior and Brain Science*, *23,* 131-46.

Landau, A.T. & Limb, C.J. (2017). The neuroscience of improvisation. *Music Educators Journal*, 27-33.

Lange, M. (1993). Lawlikeness. *Nous*, *27*, 1-21.

Laske, O. (1977). *Music, memory, and thought: Explorations in cognitive musicology.* Ann Arbor, MI: University Micro Films International.

Leman, M. (2007). *Embodied music cognition and mediation technology.* Cambridge, MA: MIT Press.

Leman, M. (2016). *The expressive moment: How interaction (with music) shapes human empowerment*. Cambridge, MA: MIT Press.

Lerdahl, F, & Jackendoff, R. (1983). *A generative theory of tonal music.* MIT Press.

Lewis, G.E. (1996). Improvised music after 1950: Afrological and eurological perspectives. *Black Music Research Journal*, 16, 91-122.

Lewis, G.E., & Piekut, B. (2016). *The Oxford handbook of critical improvisation studies, Vol. 1.* Oxford U. Press.

Liang, S., Wu, X., & Jin, F. (2018). Gut-brain psychology: Rethinking psychology from the microbiota-gut-brain axis. *Frontiers in Integrative Neuroscience*, *12*, 33.

Limb, C.J., & Braun, A.R. (2008). Neural substrates of spontaneous musical performance: An fMRI study of jazz improvisation. *PLoS ONE, 3*, e1679.

Linson, A., & Clarke, E.F. (2017). Distributed cognition, ecological theory and group improvisation. In E.F. Clarke & M. Doffman (Eds.), *Disembodied creativity: Collaboration and improvisation in contemporary music* (pp. 52-69). Oxford U. Press.

Linson, A., Clark, A., Ramamoorthy, S., & Friston, K. (2018). The active inference approach to ecological perception: general information dynamics for natural and artificial embodied cognition. *Frontiers in Robotics and AI*, *5*, 1-22.

Loehr, J.D., Kourtis, D., Vesper, C., Sebanz, N., & Knoblich, G. (2013). Monitoring individual and joint action outcomes in duet music performance. *Journal of Cognitive Neuroscience, 25*, 1049–1061.

Longuet-Higgins, H.C. (1979). The perception of music. *Proceedings of the Royal Society, B 205*, 307-322.

Longuet-Higgins, H.C. (1994). Artificial intelligence and musical cognition. *Philosophical Transactions of the Royal Society, A, 349*, 103-113.

Love, S.C. (2016). The jazz solo as virtuous act. *The Journal of Aesthetics and Art Criticism, 74*, 61-74.

Love, S.C. (2017). An ecological description of jazz improvisation. *Psychomusicology: Music, Mind, and Brain, 27*: 31-44.

Lutterbie, J. (2011). *Toward a general theory of acting: Cognitive science and performance.* Palgrave MacMillan.

Martin, K.A.C. (1994). A brief history of the "feature detector." *Cerebral Cortex*, *4*, 1-7.

MacDonald, R.A.R., & Wilson, G.B. (2006). Constructions of jazz: how jazz musicians present their collaborative musical practice. *Musicae Scientiae, 10*, 59-83.

McDowell, J. (2008). Some remarks on intention in action. *The Amherst Lecture in Philosophy* Lecture 6.

Menin, D., & Schiavio, A. (2012). Rethinking musical affordances. *Avant, III*, 202-215.

Meyer, L.B. (1956). *Emotion and meaning in music*. U. Chicago Press.

Michael, J. (2017). Music performance as joint action. In M. Lesaffre, P-J. Maes, & M. Leman (Eds.), *The Routledge companion to embodied music interaction* (pp. 160-166). Routledge.

Michael, J., Sebanz, N., & Knoblich, G. (2016). The sense of commitment: A minimal approach. *Frontiers in Psychology, 6*, 1-11.

Michael, J., & Wolf, T. (2014). Why apply a hierarchical predictive processing framework to music perception and performance. *Empirical Musicology Review*, *9*, 177-182.

Michaels, C.F., & Palatinus, Z. (2011). A ten commandments for ecological psychology. In L. Shapiro (Ed.), *The Routledge handbook of embodied cognition* (pp. 19-28). Routledge.

Mithen, S. (2005). *The singing neanderthals: The origins of music, language, mind, and body.* Harvard U. Press.

Moran. N. (2014). Social implications arise in embodied music cognition research which can counter musicological 'individualism'. *Frontiers in Psychology, 5*, 1-10.

Moran N., Hadley L.V., Bader M., & Keller, P.E. (2015). Perception of 'back-channeling' nonverbal feedback in musical duo improvisation. *PLOS ONE,* 10: e0130070.

Mylopoulos M., & Pacherie, E. (2018). Intentions: The dynamic hierarchical model revisited. *WIREs Cognitive Science,* e1481.

Namba, S. & Kuwano, S. (2004). Environmental acoustics: Psychological assessment of noise. In J.G. Neuhoff (Ed.), *Ecological acoustics* (pp. 176-188). San Diego: Elsevier.

Narmour, E. (1990). *The analysis and cognition of basic melodic structures: The implication realization model*. Chicago: U. of Chicago Press.

Nöe, A. (2004). *Action in perception*. Cambridge, MA: MIT Press.

National Public Radio (2000). Body and Soul. *National Public Radio*, Online, Accessed on 11/7/2019 at: https://www.npr.org/2000/03/06/1071243/body-and-soul

Nussbaum, C.O. (2007). *The musical representation: Meaning, ontology, and emotion*. Cambridge, MA: MIT Press.

Oechslin, M.S., Van De Ville, D., Lazeyras, F., Hauert, C., & James, C.E. (2013). Degree of musical expertise modulates higher order brain functioning. *Cerebral Cortex*, *23*, 2213-2224.

Orlandi, N. (2018). Predictive perceptual systems. *Synthese, 195*, 2367-2386.

Overy, K. & Molnar-Szakacs, I. (2009). Being together in time: Musical experience and the mirror neuron system. *Music Perception. 26*: 489-504.

Pacherie, E. (2008). The phenomenology of action: A conceptual framework. *Cognition, 107*, 179-217.

Pacherie, E. (2012). The phenomenology of joint action: Self-agency vs. joint-agency. In A. Seemann (Ed.), *Joint attention: New developments* (pp. 343-389). Cambridge, MA: MIT Press.

Pacherie, E. (2018). Motor intentionality. In A. Newen, L. de Bruin & S. Gallagher (Eds), *The Oxford handbook of 4e cognition* (pp. 369-388). Oxford: Oxford U. Press.

Pacherie., E. & Dokic, J. (2006). From mirror neurons to joint actions. *Journal of Cognitive Systems Research, 7*, 101-112.

Pachet, F., Roy, P., & Foulon, R. (2017). Do jazz improvisers really interact? The score effect in collective jazz improvisation. In M. Lesaffre, P-J. Maes, & M. Leman (Eds.), *The Routledge companion to embodied music interaction* (pp. 167 – 176). Routledge.

Pearce, M., & Rohrmeier, M. (2012). Music cognition and the cognitive sciences. *Topics in Cognitive Science, 4*: 468-483.

Pfordresher, P.Q. (2008). Auditory feedback in music performance: The role of transition-based similarity. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 708-725.

Pfordresher, P.Q. & Palmer, C. (2002). Effects of delayed auditory feedback on timing of music performance. *Psychological Research, 16*, 71-79.

Pinker, S. (1997). *How the mind works*. W.W. Norton & Company

Poort, J., Raudies, F., Wannig, A., Lamme, V.A.F., & Neumann, H. (2012). The role of attention in figure-ground segregation in areas V1 and V4 of the visual cortex. *Neuron, 75*, 143-156.

Pressing, J. (1988). Improvisation: Methods and a model. In J.A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition* (pp. 129–178). New York, NY: Oxford U. Press.

Pressing, J. (1998). Psychological constraints on improvisational expertise and communication. In B. Nettl & M. Russel (Eds.), *In the course of performance: Studies in the world of musical improvisation* (pp.47-68). U. Chicago Press.

Preston, B. (2012). *A philosophy of material culture: Action, function, and mind*. Routledge.

Quinton, A. (1975). Social objects. *Proceedings of the Aristotelian Society, 76*, 1-27.

Raffman, D. (1993). *Language, music, and mind*. Cambridge, MA: MIT Press.

Raja, V. (2018). A theory of resonance: Towards an ecological cognitive architecture. *Minds & Machines, 28*, 29-51.

Ramstead, M.J.D., Veissière, S.P.L., & Kirmayer, L.J. (2016). Cultural affordances: Scaffolding local worlds through shared intentionality and regimes of attention. *Frontiers in Psychology, 7*, 1-21.

Ramstead, M.J.D., Kirchoff, M.D., Constant, A., & Friston, K. (2019). Multiscale integration: beyond internalism and externalism. *Synthese*, Online Version https://doi.org/10.1007/s11229-019-02115-x.

Rauss, K., & Pourtois, G. (2013). What is bottom-up and what is top-down in predictive coding? *Frontiers in Psychology, 4*, 1-8.

Reed, E.S. (1996). *Encountering the world. Toward an ecological psychology*. New York, NY: Oxford U. Press.

Reybrouck, M. (2012). Musical sense-making and the concept of affordance: an ecosemiotic and experiential approach *Biosemiotics, 5*, 391-409.

Rietveld, E., & Kiverstein, J. (2014). A rich landscape of affordances. *Ecological Psychology, 26*, 325-352.

Rietveld, E., Denys, D., & Van Westen, M. (2018). Ecological-enactive cognition as engaging with a field of relevant affordances: The skilled intentionality framework (SIF). In A. Newen, L. de Bruin, & S. Gallagher (Eds.), *Oxford handbook of 4e cognition* (pp. 41-70). Oxford: Oxford U. Press.

Risjord, M. (2014). Structure, agency, and improvisation. In J. Zahle & F. Collin (Eds.), *Rethinking the individualism-holism debate* (pp. 219-236). Synthese Library 372.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research, 3*, 131-141.

Ryan, K., & Schiavio, A. (2019). Extended musicking, extended mind, extended agency. Notes on the third wave. *New Ideas in Psychology*, *55*, 8-17.

Ryle, G. (1949). *The concept of mind*. U. Chicago Press.

Ryle, G. (1976). Improvisation. *Mind, 85*, 69-83.

Schaefer, R. (2014). Mental representations in musical processing and their role in action perception loops. *Empirical Musicology Review, 9*, 161-176.

Schiavio, A. (2015). Action, enaction, and inter(en)action. *Empirical Musicology Review, 9*, 254-262.

Schiavio, A., & Høffding, S. (2015). Playing together without communicating? A pre-reflective and enactive account of joint musical performance. *Musicae Scientiae*, 1-23.

Schmitz, L., Vesper, C., Sebanz, N., & Knoblich, G. (2018). Co-actors represent the order of each other's actions. *Cognition, 181*, 65-79.

Searle, J. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge, UK: Cambridge U. Press.

Searle, J. (1990). Collective intentions and actions. In P.R. Cohen, J. Morgan and M. Pollack (Eds.), *Intentions in communication* (pp. 401-415). Cambridge, MA: MIT Press.

Searle, J. (2014). *Making the social world: The structure of human civilization.* Oxford, UK: Oxford U. Press.

Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others' actions: just like one's own? *Cognition, 88,* B11-B21.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in Cognitive Science, 10,* 70-76.

Sebanz, N., & Knoblich, G. (2009). Prediction in joint action: What, when, and where. *Topics in Cognitive Science, 1,* 353-367.

Shapiro, L. (2010). *Embodied cognition.* (1st ed.). New York: Routledge.

Shepard, R.N. (1984). Ecological constraints on internal representation: Resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychological Review, 91,* 417-447.

Small, C. (1998). *Musicking: The meanings of performance and listening.* Wesleyan U. Press.

Srinivasan, M.V., Laughlin, S.B., & Dubs, A. (1982). Predictive coding: a fresh view of inhibition in the retina. *Proceedings of the Royal Society of London, Series B, Biological Sciences, 216,* 427–459.

Stoffregen, T.A. (2003). Affordances as properties of the animal-environment system. *Ecological Psychology, 15,* 115-134.

Sudnow, D. (2001). *Ways of the hand: A rewritten account*. Cambridge, MA: MIT Press.

Sutton, J., Harris, C.B., Keil, P.G., & Barnier, A.J. (2010). The psychology of memory, extended cognition, and socially distributed remembering. *Phenomenology and the Cognitive Sciences, 9,* 521-560.

Taylor, A. (1993). *Notes and tones: Musician-to-musician interviews.* Da Capo Press.

Theiner. G., Allen, C., & Goldstone, R.L. (2011). Recognizing group cognition. *Cognitive Systems Research, 11,* 378-395.

Tollefsen, D. (2005). Let's pretend! Children and joint action. *Philosophy of the Social Sciences, 35,* 75-97.

Tollefsen, D. (2014). A dynamic theory of shared intention and the phenomenology of joint action. In S.R. Chant, F. Hindriks, and G. Preyer (Eds.), *From individual to collective intentionality: New essays* (pp.13-33). Oxford U. Press.

Torrance, S., & Schumann, F. (2019). The spur of the moment: What jazz improvisation tells cognitive science. *AI & Society, 34*, 251-268.

Treisman, A. (1999). Solutions to the binding problem progress through controversy and convergence. *Neuron, 24*, 105-125.

Tuomela, R. (2007). *The philosophy of sociality: The shared point of view*. Oxford U. Press.

Turvey, M.T. (1992). Affordances and prospective control: An outline of the 0ntology *Ecological Psychology, 4*, 173-187.

Turvey, M.A., Shaw, R.E., Reed, E.S., & Mace, W.M. (1981). Ecological laws of perceiving and acting: In reply to Fodor and Pylyshin (1981). *Cognition,* 237-304.

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 & Science, 1,* 1-18.

Van Djik, L., & Rietveld, E. (2018). Situated anticipation. *Synthese,* Online Version, https://doi.org/10.1007/s11229-018-02013-8.

Veissière, S.P.L, Constant, A., Ramstead, M.J.D., Friston, K.J., & Kirmayer, L.J. (Forthcoming). Thinking through other minds: A variational approach to cognition and culture. *Behavioral and Brain Sciences,* Page Citations from doi: https://doi.org/10.1017/S0140525X19001213.

Vesper, C., Abramova, E, Bütepage, J., Ciardo, F., Crossey, B., Effenberg, A., Hristova, D.,… & Wahn, B. (2017). Joint action: Mental representations, shared information, and general mechanisms for coordinating with others. *Frontiers in Psychology, 7*, 1-7.

Vetter, P., Edwards, G., & Muckli, L. (2012). Transfer of predictive signals across saccades. *Frontiers in Psychology, 3*, 176.

Walton, A., Richardson, M.J., Langland-Hassan, P., & Chemero, A. (2015). Improvisation and the self-organization of multiple musical bodies. *Frontiers in Psychology*: https://doi.org/10.3389/fpsyg.2015.00313.

Walton, A., Richardson, M. J., Langland-Hassan, P., Chemero, A., & Washburn, A. (2018). Creating time: Affording social collaboration in music improvisation. *Topics in Cognitive Science,* 95-119.

Warren, W.H. (1984). Perceiving affordances: Visual guidance of stair climbing. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 683-703.

Wilson, G.B., & MacDonald, R.A.R. (2017). The construction of meaning within free improvising groups: A qualitative psychological investigation. *Psychology of Aesthetics, Creativity, and the Arts, 11*, 136-146.

Windsor, W.L. & de Balzac, C. (2012). Music and affordances. *Musicae Scientiae, 16*, 102-120.

Wolf, T., Sebanz, N., & Knoblich, G. (2018). Joint action coordination in expert-novice pairs: Can experts predict novices' suboptimal timing? *Cognition, 178*, 103-108.