PREPRINT COPY: Published chapter can be found here https://link.springer.com/chapter/10.1007/978-3-030-72116-9_1.

Sociocultural and design perspectives on AI-based music production: why do we make music and what changes if AI makes it for us?

Oliver Bown
Interactive Media Lab, Faculty of Art & Design
UNSW, Paddington, NSW, 2021
o.bown@unsw.edu.au

May 26, 2020

Introduction

The recent advance of artificial intelligence (AI) technologies that can generate musical material (e.g., [18, 19, 75, 37]) has driven a wave of interest in applications, creative works and commercial enterprises that employ AI in music creation. Most fundamental to these endeavours is research focused on the question of how to create better algorithms that are capable of generating music, but the wider issue of application and use draws on a diverse range of fields including design, psychology, creative practice and sociocultural factors. This chapter takes a closer look at the sociological and design dimensions of AI music generation, and offers researchers in AI and music a number of themes and references that can be used to help frame creative practice and applications, supporting the design and evaluation of creative music systems.

I begin by considering the more practice-based origins of AI-based music generation, which I suggest has traditionally taken a more philosophical orientation towards the question of musical intelligence and creative autonomy in machines. Then I consider how design has given the field a more functional applied focus which has to some extent drawn attention away from the philosophical concerns of earlier creative practitioners. Then I consider how sociological views, which remain relatively marginal to the practice of AI music, are beginning to influence the field and contribute a much-needed vantage point for understanding human musical behaviour.

The Philosophical Era

In the 1950s Lejaren Hiller and Leonard Isaacson composed the *Illiac Suite* [46], widely recognised as the first score composed with the creative input of algorithms running on a digital computer, the ILLIAC 1 computer, based at the University of Illinois at Urbana-Champaign. The *Illiac Suite* consists of four movements, each of which explored different methods of algorithmic composition including hand-coded rules and the now widely used stochastic generation method of Markov modelling. It is well over half a century since this first experiment in computer music generation, and the capabilities of computers have advanced unimaginably. Subfields of Computer Science have made great advances, particularly over the last decade in Deep Learning [71] (a summary of Deep Learning applications in music can be found here [18, 19]). Programs far more complex than Hiller and Isaacson's run in realtime on $10 credit card sized computers, and school children are learning to code using programming languages that abstract away the complexities of the machine code they would have had to use. Yet listening to the *Illiac Suite*, you might think that the authors had solved automated music generation on the very first attempt. The music sounds harmonious and richly-structured, original and emotional. You might conclude upon listening that automated music generation is not a hard problem.

In fact, on the contrary, this field possesses some of the characteristics of a *wicked problem* [22], that is ill-defined and potentially unsolvable due to its complexity and sociocultural embeddedness. Relevant features of wicked problems include that there is no definitive formulation, no ultimate test and no stopping rule. Researchers are definitely still trying to discover functional music generation algorithms today, that could effectively generate new and original musical works from the ground up. Unsurprisingly, this disconnect between us thinking of the *Illiac Suite* as a success and yet still finding automated music generation a hard problem arises because of the complex context of creative production.

Hiller and Isaacson were composers (amongst other things), and their approach to the *Illiac Suite* took a form that continued for the rest of the century to be, arguably, the predominant approach to computer-generated music experiments; creative practitioners with hacking skills (or with paid programming support) programming computers in tasks of algorithmic composition (there are significant exceptions, of course, like the creation of more traditional end-user tools such as Microsoft's *Songsmith*). Here, the creative interaction between human and software system takes the form of an iterative cycle of programming and reviewing machine output, followed by any further processes the human author wishes to add to complete the work (arranging, orchestrating and performing the work, for example). The process is summarised in a simplified form in Figure 1.

This diagram purposefully highlights the limited scope of the machine and the various ways in which creative authority is maintained by the human composer. For example, if the system produces outputs that don't sound right the first time, the composer can modify the program code or parameters through trial and error – a form of search enacted by the composer. They can cherry-pick outputs, and they can perform additional work on the resulting outputs, modifying them or engaging in additional steps such as arranging, orchestrating or rendering the result.

Thus conceived, the challenge of making better algorithmic composition systems could be defined in terms of moving more of these stages and decisions into the computer's scope, to the point where a computer is more autonomous in the creation of music, or alternatively making the computer perform so well in the given role that it minimises the need for the composer to iteratively search for the 'right' output. Historically, this has been the focus of much academic work in computer generated music,

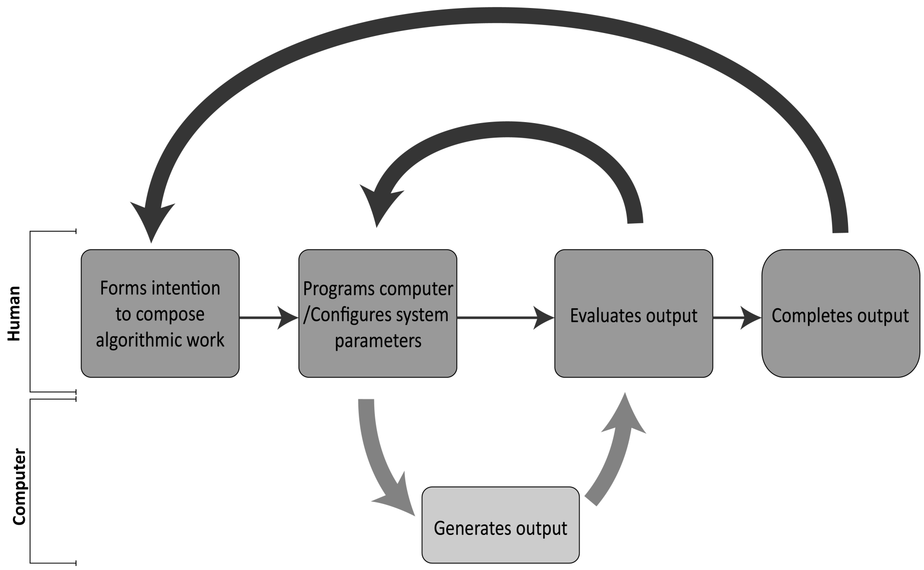

Figure 1: Typical human-computer creative interaction in algorithmic composition where the composer is either the programmer or user of an existing program.

underpinned by a vision of creating machines that fulfil the role of human-equivalent artists. Practitioners of algorithmic composition have endeavoured to create systems that allow them to be as hands-off as possible, intending to hand over autonomy and control to the machine in order to fulfil this vision of the machine performing the artist's job. We can think of this in terms of timescales or levels of creative agency. At a very low level we might think of the creative work that goes into the choice of a single note in a sequence. Higher up we might think of the composition of an individual melody. Up and up at the other extreme we might think of the kind of creative development a practitioner engages in over their life time involving choices of styles, paradigms and strategies; in theory these might all be creative processes that could be codified in computer programs.

Such a vision was popularised by writers such as Ray Kurzweil [51] who prophesised that autonomous machine artists were imminent and, as in other areas of AI, would begin to take their place alongside human artists. For many artists, this endeavour to experiment with creative autonomy in machines was primarily grounded in an aesthetic philosophical interest; the creative work acting both as a simulation experiment as well as a kind of cultural probe [41], stimulating reactions. David Cope [27, 28] famously upset musicologists with machine generated compositions that he disguised as human works. He viewed his works as a proof-of-concept that musical composition could be subjected to mechanical instructions, and that much of the defensiveness found in responses to his work was hubristic and based on an essentialist view that human creativity was profound and inimitable. In the parallel world of generative visual art systems, pioneer Harold Cohen [25] used his work, co-creating with rule-based software of his own making, to explore how we attribute authorial agency to machines:

> "If a photographer takes a picture, we do not say that the picture has been made by the camera. If, on the other hand, a man writes a chess-playing program for a computer, and then loses to it, we do not consider it unreasonable to say that he has been beaten by the computer". [25]

But since these composer-centric efforts in algorithmic composition are always ultimately (and heavily – despite various claims made) in the creative control of the practitioner, it becomes tricky to define quite how the system is able to take on creative autonomy, no matter how good the results are and how little, metaphorically speaking, the human artist's hand touched the canvas. For example, Brian Eno's celebrated generative music experiments [38] were purposefully musically simple, constrained to a pleasant melodic space not dissimilar to the output of aeolian harps. The *Illiac Suite*'s fourth movement sounds significantly more abrasive and challenging than the other movements, reflecting the weirder results of their Markov experiments. In both cases, we are hearing the aesthetics that emerge when a composer grapples with algorithmic possibilities; arguably less important is the detail filled in by the algorithm's decisions. It is possible that the trickiness of unpacking and evaluating such efforts–the wickedness of the problem–has, over time, whittled away this once strong interest in making autonomous machine artists.

Moreover, there is a wider context of engagement in the sociocultural environment that frames such cultural production; the human arrives at this interaction having formed an intention to do so, usually with a specific style, agenda and audience in mind, and, importantly, motivated by social objectives such as engagement with a community or social status. They may also exert other forms of influence on the sociocultural environment; for example, if they are successful they may be more empowered to influence what constitutes good music. If all of this culturally grounded activity is part of being an artist, then the full hand-over of being-an-artist to a machine becomes unimaginably complex and beyond the scope of anyone's capability.

Friction between engineering and sociologically-grounded perspectives arises here: from an engineer's perspective, where simplifying first-approximations to problems are used purposefully to scaffold more substantial solutions, it would seems reasonable to separate out such questions of social context from questions of what note goes next or how to arrange a composition for dramatic effect, and the success of music generation tools at mimicking musical style and proposing interesting new content is testament to this approach. However, from a human-scientist's perspective, the engineer's simplification may be untenable; the result of their work is a strange beast that does not correspond exactly with anything in the pre-existing world. It is as much a novel hybrid than it is a model of any particular aspect of human behaviour. This is an important difference of perspective that can be hard to fully appreciate.

Creative Cognition and Lofty versus Lowly Computational Creativity

In the field of Computational Creativity [23], a common tenet is that the mere generation of creative content, although a critical subcomponent, is far from a complete model of human creative production [78]. An act of creative production, it is held, requires evaluation and iteration of outputs; when humans make music, art and other forms of cultural production, a key component of their process is reflection and iteration. Algorithms that do not do this are deficient and, it could be said, have no sense of the aesthetic and cultural world they act in.

In practical terms, the majority of current generative arts practices are situated in the domain of ”mere generation”. This includes most of the work of creative coding artists from Harold Cohen and David Cope onwards, who shepherd their generative systems, but it also includes the bulk of advanced machine learning (ML) systems applied to music that perform predictive modelling of what should come next in a sequence, not involving any creative iteration or evaluation. These ’lowly’ forms of generative practice may generate creative works of stunning originality and sophistication, but they have no meta-awareness they are doing so. They are not culturally aware agents and they have no sense of judgement or inherent taste themselves. The ’lofty’ goal of Computational Creativity is to create systems that perform genuine human-like creative cognition, which involves a host of cognitive processes and is largely understood as a search process.

The Design Turn

In recent years, design thinking and design practices have become increasingly foregrounded in the study of the application of AI to music (e.g., [73, 15, 39]). This is in step with design's more general rise to prominence, especially in technology innovation where design methods now enable a richer and more productive cycle of feedback between the work of the engineer and the experience and needs of the human user. It has also arisen naturally within the music AI field in recognition of the importance of interfaces and user-interactions even when the ultimate goal is to build an ’autonomous’ intelligent system. Everything has an interface, even a sophisticated robot that can converse in natural language, or a realtime improvisational musical agent that interacts via musical sound. As AI-powered systems emerge in multiple different fields, it has become increasingly evident that the design of interactions occurs all the way up the scale of AI sophistication. Even if you are conversing with a machine in plain English, its behaviour is packed with design decisions, from the overall interface and interaction context to the specific responses and behavioural traits. Thus, design considerations clearly extend beyond the graphical and tangible aspects of an interface; non-graphical voice-audio or music-audio interfaces present their own unique challenges for interaction design. For example, the makers of realtime improvisational interfaces must still grapple with designed interactions such as turn-taking strategies or ways for the system to signal to the user that it is responding, whether it is hearing properly, and so on (e.g., as discussed in [20]).

Although they rarely framed their discussion in terms of interaction design, those early innovators such as Hiller and Isaacson, Cope and, in the visual domain, Harold Cohen, could all be seen to be interacting with their systems in a variety of ways, even if through the code interface itself, engaging in heuristic creative strategies, evaluating and iterating. In doing so their work inevitably highlights many interesting interaction design concepts. Today design thinking horizontally penetrates other fields from business to computer science, with music AI systems as natural an area of application as any other; current creators of AI music systems are more likely to apply design terminology and methods, following the leading work by researchers such as Rebecca Fiebrink [39], Thor Magnusson [55], Nick Collins [26], Andrew Brown [20] and François Pachet [3] (also see Anna Kantosalo's work in non-music domains [49]). Areas of practice such as "live coding" have focused more explicitly on the design of usable and creatively productive programming languages and environments, work that has extended into the design of usable generative and AI-based systems. This work draws on research into the design of creatively empowering and productive software systems, such as Blackwell's work on the cognitive dimensions of notations [10], Shneiderman, Resnick, Candy, Edmonds and colleagues' work on creativity support tools [72, 66, 6], and more general usability principles such as Nielsen's heuristics for user interface design [60] and key interaction design work such as that of Donald Norman [61].

Thus increasingly, the question of how good interaction design can support the creation of successful AI music tools has become more commonly posed (for example [26, 73, 14, 55, 39]), and a general shift can be seen away from the more philosophical concerns of autonomous music composition agents, towards AI tools that support creative production. This is in part due to the maturation of the technology and the potential for business opportunities that invite a user-centered design perspective to solve the problem of making usable technology.

The design of usable AI music systems can be divided very roughly into two categories: tools targeting creators, those people already involved in the production of music; and tools targeting non-musician end-users who need music services, including professionals such as film producers, their equivalent amateur or prosumer creators (such as amateur YouTube contributors), and other types of music consumers. In both cases, as with other areas of technology innovation, the use-cases to which new AI music technologies are applied can be broadly divided into those use-cases that effectively model an existing scenario, and those use-cases that are novel.

Regarding tools for creators, there are a number of common creative practice contexts and activities that point to some of the main uses for AI in music. Some major areas that cover a great deal of the scope of creator-based AI music production are as follows:

- Supporting a composition task in an offline context (i.e., not performed live), for example when working in a Digital Audio Workstation, on a paper score, or composing with an instrument.

- Creating, configuring, and performing with intelligent performance systems such as improvising musical agents and intelligent augmented instruments, where the system is perceived as having creative agency in some capacity. This has been a popular topic of discussion in communities such as the Live Algorithms for Music and Musical Metacreation research groups.

- Meta-designing musical interfaces to be used in creation or performance. This area is exemplified by the work of Rebecca Fiebrink with her Wekinator software, which can be used for rapid on-the-fly mapping by non-programmer users in a range of contexts [39].
- Developing bespoke systems with niche applications. A recent popular example is Holly Herndon's AI choir, which does not aim for a dominant creative role by AI in the musical composition, but a more specific creative function within her compositional work.

The first area is vast and I will consider it in a bit more detail. AI supported composition may be a collaborative or solo activity in any number of musical cultural niches from commercial studio production to highly experimental art music to educational contexts. Common tasks include melodic-, chord- or drum-pattern creation, arrangement, orchestration, harmonisation, timbre selection, expressive performance rendering (as in "groove quantise") and related production tasks such as mixing and mastering (as performed by the successful commercial software service *Landr* [53]). Contextual factors for the task may include adaption to existing musical content (e.g., selecting a chord progression to fit a melodic line), and selecting styles and other parameters (e.g., achieved through selecting a training corpus). Two important use-case paradigms here are: rapid ideation, as performed by *Aiva* [1], where the system's main purpose is to support the rapid ideation of potentials; and a producer-session-musician relationship, where the system's main purpose is to provide modular adaptive units for insertion into a composition, as popularised in amateur production by Apple's *Garage Band* software [2]. Slightly distinct from these two paradigms is the situation in which there is more of a conceptual commitment to the idea of the system being an autonomous independent creator, or more loosely sometimes acting as a 'collaborator' (examples include the Sony CSL-produced *Daddy's Car* track [32]). One other related category is the creation of adaptive compositions such as music for games or VR/AR experiences, where the preparation of a composition is performed offline, but its ultimate structuring occurs at listening time. However, this category of activity is more correctly placed in the "Tools for Consumers" category.

Tools for consumers might include systems for democratising music production and generative music for games and other interactive experiences. We can imagine that as such tools emerge the specific use-cases in which consumer-based generative music might occur will evolve. For example, with the rise of predictive tools in text messaging and richer text media such as emojis and gifs we can imagine the automation of creative content such as images or music becoming applied to rapid-fire social interactions.

Design evaluation

The design process is strongly focused on evaluation of tools in the hands of users, feeding back into improved designs in an iterative cycle. For example, tools to support music ideation can be evaluated in terms of the speed of development of new musical ideas, how well they enable a creator to break a creative block or discover new directions (as reported by the creator, for example), or how well they augment someone's musical skill (for example, could a system that enables style transfer support an expert jazz musician to transfer their skills to latin music styles?).

Compared to the earlier philosophical perspective, the system's creative autonomy becomes a broader but somewhat more measurable issue from a design perspective:

- Does the system do music well (canonically) under given stylistic or task constraints? In many contexts, such as generating a piece of music for a YouTube video, we aren't necessarily concerned with whether a system is innovative, but whether it can produce functional music that conforms to stylistic requirements. Thus a need for typicality may be key [68].
- Does the system creatively innovate? Does it evoke a sense of being talented, maverick, surprising or inspiring? This may be more useful in the context of breaking creative block.
- Does the system give the *impression* of either performative agency (demonstrating autonomy in a performance context) or memetic agency (demonstrating agency over a cultural timescale) [16]? Giving the impression of agency, regardless of what is actually produced, may be an important factor in certain contexts such as interactive performance contexts where liveness is at stake.

Here, the context of the creativity activity has a nuanced impact on how we might judge the importance of any system's intelligence and autonomy. In a performance context, for example, the sense of autonomy may be an important aspect of the experience of the work. In a creative production context, the system's creative autonomy could potentially problematise the author's sense of ownership of the work, if there is a strong relationship between authorship and identity. In a commercial context, this may be less the case if the identity of the work is more anonymous, but creators may nevertheless be concerned about the technicalities of ownership, i.e., issues of copyright.

Such concerns regarding contextual factors in the evaluation of music AI systems, from a design perspective, naturally invite a number of sociological questions, which form the next stage in this discussion.

The Sociological View

A more recent development that builds in turn on design practices is the introduction of sociological perspectives on how we think about and study creative practices in the AI music space. Sociological perspectives naturally complement design, contributing practical extensions to design research with ethnographic methods for the observation of producers or consumers of music. Since designers focus on understanding use in practice, social and cultural factors are naturally important; how we use cars or kitchens, for example, and therefore how they should be designed, is deeply entangled with our social patterns of behaviour, which vary greatly across cultural contexts. Ethnography emphasises situated qualitative research that holistically takes into account the cultural context in which people operate, where 'culture' might refer equally to what differs between bedroom techno producers and commercial pop producers, as to what differs between religious rites in Cuban Catholic and Thai Buddhist groups. Music can safely be considered exemplary of a cultural activity; we inhabit niches of musical style specific to our cultural backgrounds, and we find music playing a role in a wide number of critical social activities from funerals and weddings to the formation of friendships and the playing out of everyday shared experiences. Yet a sociological perspective has revealed how even seemingly non-cultural activities such as scientific research are shaped in very important ways by cultural factors (e.g., [50, 54]. In this case, whilst the principle of scientific objectivity may be closely adhered to, science is still carried out by individuals with personal ambitions, biases and social relations, and within systems of social organisation that strongly influence behaviour.

A sociological perspective also provides a frame, building on relevant social theory, with which to conceptualise AI music activity. It is not a neutral method but a creative one, where conceptual frameworks can guide how research is done. One of the pioneers of Sociology and Anthropology, Emile Durkheim [36], believed that a social perspective is to psychology what chemistry is to physics: the phenomena of the social are irreducible to explanation at the level of individual human psychology. Durkheim developed the concept of the collective consciousness to capture the notion that societies operate in systems of share beliefs, ideas, habits, styles and attitudes.

In modern work on music, Born [12], for example, considers how social relations – such as those amongst musical collaborators, between themselves and their audiences, or in the wider social environment as reflected upon by musical artists – influence and are influenced by musical practices, styles and cultures in complex but traceable ways that are subject to useful analysis. A band's social organisation may reflect its politics, for example.

Cluster concepts and emic versus etic definitions

The sociological perspective is valuable for a number of reasons but one of the most fundamental is how it enables us to distinguish the analysis of the practice, as seen from 'outside', from the perspective that emerges from the practice itself – this is something that distinguishes it from art theory or musicology, for example. In Anthropology, an emic perspective is a perspective from within the system being observed and an etic perspective is that of the (in theory) independent observer. This distinction is immediately enlightening when we consider how we might define two important but ambiguous terms: creativity and art. From an emic perspective, the use of both terms may be observed in practice. For example, we might notice that creativity is closely associated by people with expression of individuality in some contexts, and more closely associated with occasional genius and a high degree of training in others (Anna Jordanous [47] was the first to bring this topic to the fore in the field of Computational Creativity). Likewise, the familiar provocations of conceptual artists to push the boundaries of what is considered art, and the predictable cries of "this is not art" from some sectors of the population, highlight how there is ongoing movement around that concept. We might therefore note that even when more formal definitions are in play, they are always contestable and blurry at the edges.

This tallies with how psychologists have come to understand how concepts are formed and used by people [59]. Rather than adhere to definitions, in practice we understand concepts in relation to *prototypes*. Things that are very similar to those prototypes are more likely to fit the concept, and their salient properties more likely to be treated as important to the concept: thus we recognise fluffy dogs as more dog-like than bald dogs. Pablo Picasso and Albert Einstein might serve as prototypical ”creative” people.

Meanwhile, we have to be able to understand what *we* (as researchers) mean by these terms even when their meaning might change from context to context. Although one solution is to seek an operationalisable definition of such terms, as has been attempted in various fields, it is arguably more useful to seek useful technical concepts that inspire an underlying social or psychological phenomenon. Examples include work by researchers such as Dissanayake (art as ”making special” [35]), Blacking (music as ”human organised sound” [8]), Bloom (”essentialism” as a structuring concept [11]) and Blackmore (meme theory [9]), who attempt to seek the essence of the behaviour in question.

Social perspectives on the psychology of creativity

A key concept in the study of the psychology of creativity is that creative tasks are those that inherently involve a process of search for an outcome that has not been considered before. Thus a strong focus of creative psychology research is the question of what strategies and heuristics support effective search. Wallas' early formulation of creative cognition [44], for example, identified the four stages of preparation, incubation, illumination and verification. During incubation, an unsolved problem goes on the cognitive backburner, so to speak, away from conscious attention, but is still being processed and might respond to salient input; as in the tale of Archimedes shouting ”eureka” in the bath. This process outline suggests at least two cycles of trial and error: one internal and subconscious and the other external and overtly evaluated. A substantial body of more recent work may deviate significantly from Wallas' early formulation, but shares a common framework of looking at the cognitive heuristics and methods of search.

Meanwhile, a longstanding body of theory has considered equivalent questions formulated through a social frame. In social system, many individuals performing search are contributing to a higher-level collective, parallel search system. Mathematicians trying to prove Fermat's Last Theorem, for example, are in competition with each other, but in doing so are collectively collaborating on the problem being solved faster at that collective level. Thus the same heuristics of creative search found in Psychology can also be seen operating at the social level, sometimes more overtly constructed in systems like market mechanisms, grand challenges in research, patent law, peer review and open data policies.

When viewed at the social level, it is worth noting that our ability to imitate becomes one of the most important cognitive abilities in the support of creative innovation: if we couldn't learn from each other then creative outcomes would never accumulate and there would be little benefit in individuals being creative (discussed for example in different ways by [67, 76, 9]). [67] in particular emphasise the importance of social learning to strategic behaviours and to the evolutionary process itself. It is common sense that those most successful in creative spheres must be trained, so that they can 'stand on the shoulders of giants'. Furthermore, creative activities operate in cultures of practice which create their own environments of objectives and evaluation (a phenomenon known in evolutionary theory as niche construction [52]). This is especially evident in the arts and there is evidence, discussed below, that the practice of artistic behaviours is tied up with the construction and maintenance of these cultural groups.

Most well-known to researchers in Computational Creativity, Mihaly Csikszentmihalyi's systems model of creativity [31] defines such a community as a collective realm in which creativity occurs, the creative individual being one 'owner' of a creative output, but the community of others who evaluate it and hold it up as worthy are just as important. Indeed, analysis by Schaffer in the domain of science suggests that often the attribution of a breakthrough to an individual can be exaggerated, diminishing the collective action of the community, a process of myth-making that may perform some functional role.

Whilst Csikszentmihalyi's work is more well-known to Computational Creativity researchers, one of the most extensive bodies of work studying communities of creative practice in Sociology is by Howard Becker [4]. Becker's work is less about predictive theories and more about rich description of the minutiae of concerns of such communities, prising apart the myriad different individuals and the relationship they each hold to the systems they interact with. For example he discusses the esoteric musical interests of undergraduate music students inhabiting small closed cliques of extreme experimentalism, a consequence of the intellectual learning environment they inhabit and compete within. Always, importantly, in Becker's studies, the relevant community is never a community of artists, but may include business people, lawyers, technicians, marketing people, philanthropists and so on, all contributing to the structures within which any other individual operates. The resulting portrait is one in which any particular individual is pursuing a complex multi-factored goal that might involve money, social status, political intentions, ethnicity, fame, tradition, authorial integrity, friendships, loyalties, relationships, courtship, lifestyle, quality of life, and rebellion against one's parents.

Becker's work highlights the very emergent nature of artistic creativity at this social level: artist A's intentions may be very specific to a context that is far beyond their design, and their work might have an effect that goes far beyond those intentions. Whist creativity in artistic and scientific domains has much in common, it seems fair to say that artistic domains have many more possible and divergent paths of future development.

Social theories of taste and identity

One of the most influential bodies of work in the sociological study of artistic behaviours comes from Pierre Bourdieu [13], who sought to understand amongst other things, artistic taste and its apparent role in social interaction. Bourdieu suggested that the way that taste and embodied knowledge of a cultural milieu is acquired is deeply related to social groups and their identities and boundaries. In essence, we absorb tacit cultural understanding very deeply and at such an early age that it becomes integral to our identity. It is hard to fake one's accent, which indicates one's social background, and the same is true of intimate knowledge, or lack thereof, of a cultural domain. Indeed, Bourdieu notes how some aspects of one's cultural identity can be acquired in education, learning about the arts at school, say, but some are so deeply embedded that they are really only acquired by those whose family life provides the right environment for early in-depth acquisition. These deep cultural acquisitions Bourdieu terms "habitus".

For Bourdieu, this acquisition and deep embodiment of taste, in turn influences social relations in profound ways:

> "Taste classifies, and it classifies the classifier. Social subjects, classified by their classifications, distinguish themselves by the distinctions they make, between the beautiful and the ugly, the distinguished and the vulgar, in which their position in the objective classifications is expressed or betrayed."[13]

This fits everday experience: we gravitate towards people with similar taste, or at least we find ourselves situated in filter-bubbles of cultures and tastes; we spend time in the same places or listening to the same radio stations as those whose tastes we share. This is a more general instance of the way that shared experience can forge social bonds. Related research has shown that information about a person's taste can influence others' judgements about them, including about their personalities [62].

Bourdieu also introduces the idea of "cultural capital". Like financial capital (wealth), this is capital accumulated by a person as embodied in tacit or explicit knowledge, style or status. Cultural capital includes one's knowledge, acquired through education, but also one's taste. By constructing this correspondence with financial capital, Bourdieu is able to show how the cultural qualities of people serve goals and tangible outcomes, and also how the properties of a cultural group define the context in which each individual must compete. The implication that being a consumer of art or music is a competitive activity was radical at the time and still sounds counter-intuitive or at least confrontational to a view of art and music as bringing people together or being of unquestioned benefit to humanity. Bourdieu portrays a more combative and politically charged world of music production and consumption. He says specifically of music; "nothing more clearly affirms one's 'class', nothing more infallibly classifies, than tastes in music" [13].

Such strands of research feed a broader potential that music and other artistic behaviours are functional, in some way, in the formation and maintenance of groups, or more generally, in the construction of individual identities and relations. Briefly, then, I consider the related field of research into music's origins, as viewed from the perspective of Evolutionary Psychology.

Why do we make and listen to music?

For some, the discussion of social functions for music must go further and seek a theoretical footing regarding why we make and listen to music. One approach is a cultural blank-slate perspective where we understand human behaviour as being so flexible as to adapt to very different cultural potentials. Imagine a child born into a richly musical culture, whether New York jazz or Javanese gamelan music or Irish folk. Through their immersion in this culture they gain a love of their local music, an expertise in it, and above all a deep appreciation of its cultural value and therefore potential personal value. Within this context they may pursue musical activities for a range of reasons. We may say that given the cultural context, the individual's behaviour can be explained in terms of the cultural motivators that have driven them to like the music and perhaps aspire to be a creator or collector of it. But a further goal of a social perspective is to understand how and why this particular cultural context emerged, not just how individuals act given the culture.

As an alternative to a blank-slate approach, we may consider a deep evolutionary perspective, in which we seek to understand how Darwinian evolutionary theory [33, 34, 57] may be used to explain how human artistic behaviours and their related social structures emerged. Darwinian evolutionary theory should be understood in the broadest sense, i.e., not simple adaptationism or genetic determinism, instead taking into account the more contemporary concepts of emergence, gene-culture coevolution and niche constructionism, amongst others. It should be noted that such a perspective is not widely popular amongst sociologists. Darwinian explanations of social phenomena were popular in the wake of Darwin's theory but became severely tainted by applications to eugenics and other morally problematic or flawed scientific initiatives.

There are, broadly speaking, four categories of theory of the evolution of human musical behaviour. In all cases, these needn't refer to genetically evolved behaviours, but be understood in terms of a more complex gene-culture coevolutionary framework.

1. **Cognitivist**: theories which posit that music supports cognitive development, helping us thinking.

2. **Consequentialist**: theories which do not posit a function for music itself, but explain musical behaviour as the consequence of a set of other evolved traits. The most well known of these is Stephen Pinker's [64] position that music is like cheesecake; it combining elements that we evolved to enjoy (like salt and sugar) with novelty and complexity, that we also evolved to enjoy, but neither cheesecake or music were themselves directly involved in our evolution.

3. **Cohesionist**: theories which posit that music is functional in binding groups, for example by creating shared rhythmic experiences that reinforce a shared identity [21, 29]. One example is Hagen and Bryant's [45] theory of music as a signalling system for a group's coalition strength, representing their ability to fight. This draws on the important handicap principle, or honest signalling theory, which states that a system of communication can evolve as long as it can't be cheated. Hagen and Bryant's theory is that because learning to perform together takes practice, it is an honest indicator of the cohesion of the group.

4. **Competitivist**: theories which posit that music evolved through competition between individuals. The most familiar example of this is the sexual selection theory of music [58], which states that males attracts females with their music performance ability, music being here an honest indicator of cognitive ability.

> But in fact the same logic that underlies sexual selection theory, based on honest signalling theory, can be applied in many other ways. Of particular interest is the potential to derive a competitivist model based on the work of Bourdieu.

In all cases, it is necessary to consider emergent factors and interactions between these different forms of explanation. Certainly, at some level, a consequentialist view must be foundational to music's evolution, since we wouldn't have music if we didn't first have a hearing apparatus, and some of the auditory perceptual abilities that clearly preceded human evolution [17]. The cohesionist and competitivist perspectives can also be seen to interact: it follows that if music plays a role in supporting cohesion within groups, then it also sets up the conditions for there to be competition within the group on the same grounds.

I only briefly touch on this subject here to highlight some of the ways in which the social nature of music generally exceeds the kinds of factors taken into account in current AI music models. Although sociologists and evolutionary psychologists generally work in different realms, both of these strands provide a footing for thinking about what underlying factors might fundamentally motivate musical enthusiasm and pleasure. This matters a great deal because it influences how we understand individual social motivations and pleasure responses related to our creation and experience of music. These factors draw attention to the great difference between what situated contexts AI systems inhabit when they create music and those that we inhabit.

Discussion

Just as the design perspective has been thoroughly embraced in AI music practice, the sociological perspective seems to be taking hold and promises to contribute to a far richer multidisciplinary field. What can AI currently do, and what can sociological and design perspectives do to both support and critique the development of AI music systems?

The above discussion points broadly to two related ways in which current state of the art ML systems do not employ 'human-like algorithms' of music creation. Firstly, most do not perform an iteration of search and evaluation. And secondly, they do not perform generative actions grounded in matters of taste, identity and human motivations.

This is not to say that we couldn't very soon derive systems that would satisfy these requirements, especially given the big data resources related to musical culture that are available. This can be achieved by training systems that are not just aware of musical content but also of the context, the cultural associations and meanings of the musical corpus being ingested by an algorithm. This requires a paradigm shift: the currently cutting-edge ML algorithms are still only predictive models (whereas in the Computational Creativity community there are in fact many examples of evaluative and iterative search-based models, such as [77, 43]). They take a corpus of music to infer expectations about what the next note will be in a sequence. Anecdotally, I have heard an ML model trained on the music of David Bowie being described as a "David Bowie simulator", yet this is far from an accurate description: the artist's awareness of the culture and his position in it is completely missing from the model – it is actually absurd to suggest you can model David Bowie, as a system, by looking at the output of that system. Models of taste, identity and motivation will become important as the next step in powering more human-like AI generation, if this is what we want.

In fact, we might position the currently most popular algorithms at the second stage of a possible four-stage hierarchy of more culturally-oriented models of taste acquisition:

1. Universal rules of aesthetics, such as those of [7], which assume that despite differences in taste, there are common rules to what makes something aesthetically pleasing;

2. Experience-specific models, sophisticated versions of which posit a universal cognitive mechanism combined with adaptation. For example, Berlyne [5] proposed models of cultural evolution in which individuals learnt and then became saturated in their preference for certain stylistic traits. Similarly, Ramachandran [65] proposes a number of existing cognitive strategies related to learning about the world stimulate aesthetic pleasure and are, equally, subject to adaptation through learning (this is another consequentialist theory). One example is the peak shift principle, which states that once we've learnt a structural property we tend to seek it in its extreme (explaining how some genres evolve towards more baroque forms);

3. More overtly curiousity-based models, in which we are stimulated by our ability to build effective models of the world, those forms we find most learnable being those we enjoy most. Versions of this concept can be found in Csikszentmihalyi's 'flow' theory [30] and in the work of Schmidthuber [70] and Wiggins (e.g., [40]);

4. Approaches in which all of the above are situated in the context of life strategies that shape an individual's interaction with a musical corpus. Arguably, we don't really have any good models of this. There are some experimental models that have tentatively explored this stage in the hierarchy, such as Saunders' curious design agents [69].

An additional level of complexity comes from the fact that musical styles and genres might embody extremely different sets of rules regarding innovation, authorship, copying, referencing, collaboration and what is considered the fixed and important parts of a composition versus those immaterial and changeable parts. Musical scores in Common Music Notation are treated in some cultures as instances of pieces of music, but many electronic music composers would not consider a score to be a suitable representation of a work, sonic timbre and microtiming being critically defined only in a machine rendering of the work. Regarding authorship, the importance of the author's identity may be far less present in commercial jingle composition than it is in hip hop. Where authorship matters, highly regarded authors are able to create value around their work. Regarding copying and innovation, in Jamaican dancehall it is common to produce many different variations using the same riddim (backing track), whereas reuse of previous material might be considered lazy in other music scenes. These types of variation impact the kinds of behaviours a human creative practitioner might engage in, cutting to the heart of how they make decisions as simple as, say, what note comes next in a melody. All of these nuances are arguably essential to marking out a music scene's identity; if current ML algorithms are to embody human-like musical behaviours, we have to ask which of these behaviours they are targeting. What type of musical role or musical activity, whether existing in the human realm or not, is the system aiming for?

We should also be conscious of how AI 'disrupts' these various cultural spheres, and recognise cultural applications as just as potentially dangerous as other areas of activity, in terms of the possible damage AI could wreak. Both the sociological and evolutionary perspectives on the potential social function of music suggest that disruption in the means of cultural production or the experience or dissemination of cultural products could, in perhaps very subtle ways, upset core processes underlying the formation and maintenance of communities. Perhaps not in the near future, but in the long term, such technologies could have devastating potential, in line with recent developments in the manipulation of electorates via social media disinformation.

Pointing out the ways in which current ML technologies are not simply comparable to human creators certainly does not not mean that such systems are not useful or effective tools. From the design stance – unencumbered by the need to argue for or aspire towards AI human-likeness, and simply focused on how such systems might enhance creativity – such tools may clearly perform a novel and potentially useful function. But I would argue that, with the exception of more practical user-interface focused work such as Fiebrink's, or very domain-specific functionality such as the Landr service, many existing music AI projects have been more conceptual than they have been practical and have yet to prove their use value in tangible examples. While it remains easy to generate hype around the use of AI in music, it remains hard to spot more sustained uses of the technology.

A sociological view can also be used to understand AI music systems as creative contributors in social processes of creativity, returning to the discussion surrounding how creators' creative authority interacts with and may give over to the contribution of an AI system (with reference to Figure 1). A system that generates completely novel musical content, even in the hands of an operator who is manipulating settings, data sets and so on, should be understood as technology that extends the basic creative mechanism of search in potentially powerful ways, as has been explored extensively by Computational Creativity researchers [63, 74, 48]. It is cognitive machinery of a sort, as seen from the perspective of Clark's extended-mind hypothesis [24]; it can be seen to possess material agency, in the language of Lambros Malafouris [56]; or secondary agency, in the language of Alfred Gell [42].

Richer analysis of such systems might fruitfully attempt to classify and categorise them in terms of what role they play in this bigger system, with a key distinction being between generators and evaluators: systems that generate output under the supervised control of an operator, and systems that analyse output in order to feed directly into a search process, or to provide information back to the human operator in a conversational interactive paradigm. Such analysis, in necessary detail, is beginning to emerge and will have a part to play in shaping the algorithmic design and the interaction design of future tools. As such systems start to be used in practice, it would seem likely that the emerging design requirements and applied objectives of such systems will move to outweigh the philosophical questions, but that new philosophical questions will come to the fore.

Acknowledgements

I thank Chloe McFadden for research assistance in the production of this chapter.

References

[1] 2020.

[2] 2020.

[3] Anna Rita Addessi and François Pachet. Young children confronting the continuator, an interactive reflective musical system. *Musicae Scientiae*, 10(1 suppl):13–39, 2006.

[4] Howard Saul Becker. *Art worlds*. Univ of California Press, 1982.

[5] D E Berlyne. *Aesthetics and Psychobiology*. Appleton-Century-Crofts, 1971.

[6] Zafer Bilda, Ernest Edmonds, and Linda Candy. Designing for creative engagement. *Design Studies*, 29(6):525–540, 2008.

[7] George David Birkhoff. *Aesthetic measure*, volume 9. Harvard University Press Cambridge, 1933.

[8] John Blacking. *Music, Culture and Experience: Selected Papers of John Blacking*. University of Chicago Press, Chicago, 1995.

[9] Susan J. Blackmore. *The Meme Machine*. OUP, New York, 1999.

[10] Alan Blackwell and Thomas Green. Notational systems–the cognitive dimensions of notations framework. *HCI Models, Theories, and Frameworks: Toward an Interdisciplinary Science. Morgan Kaufmann*, 2003.

[11] Paul Bloom et al. *How pleasure works: The new science of why we like what we like*. Random House, 2010.

[12] Georgina Born, Eric Lewis, and Will Straw. *Improvisation and social aesthetics*. Duke University Press, 2017.

[13] Pierre Bourdieu. *Distinction: A social critique of the judgement of taste*. Harvard University Press, 1984.

[14] O Bown and A R Brown. Interaction design for metacreative systems. In *New Directions in Third Wave Human-Computer Interaction: Volume 1 - Technologies*. Springer, 2018.

[15] Oliver Bown. Empirically grounding the evaluation of creative systems: Incorporating interaction design. In *Proceedings of the 5th International Conference on Computational Creativity*, Ljubljana, June 2014.

[16] Oliver Bown, Alice Eldridge, and Jon McCormack. Understanding interaction in contemporary digital music: from instruments to behavioural objects. *Organised Sound*, 14(02):188–196, 2009.

[17] A. S. Bregman. *Auditory Scene Analysis*. MIT Press, Cambridge: MA, 1990.

[18] Jean-Pierre Briot, Gaëtan Hadjeres, and Francois Pachet. Deep learning techniques for music generation. Technical report, Sorbonne Universites, UPMC Univ Paris 06, CNRS, LIP6, Paris, France, 2017.

[19] Jean-Pierre Briot and François Pachet. Music generation by deep learning-challenges and directions. *arXiv preprint arXiv:1712.04371*, 2017.

[20] Andrew Brown, Toby Gifford, and Bradley Voltz. Controlling interactive music performance (cim). In *Proceedings of the Fourth International Conference on Computational Creativity*, page 221, 2013.

[21] S. Brown. Contagious heterophony: A new theory about the origins of music. *MusicæScientiæ*, 11(1):3–26, Spring 2007.

[22] Richard Buchanan. Wicked problems in design thinking. *Design issues*, 8(2):5–21, 1992.

[23] Amílcar Cardoso, Tony Veale, and Geraint A Wiggins. Converging on the divergent: The history (and future) of the international joint workshops in computational creativity. *AI Magazine*, 30(3):15, 2009.

[24] A. Clark. *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*. Oxford University Press, 2003.

[25] Harold Cohen. Parallel to perception: Some notes on the problem of machine-generated art. *Computer Studies*, 4(3/4), 1973.

[26] Nick Collins. *Towards Autonomous Agents for Live Computer Music: Realtime Machine Listening and Interactive Music Systems*. PhD thesis, Centre for Science and Music, Faculty of Music, University of Cambridge, 2006.

[27] D. Cope. Computer modelling of musical intelligence in EMI. *Computer Music Journal*, 16(2):69–83, 1992.

[28] D. Cope. *Experiments in Musical Intelligence*. A-R Editions, Madison, WI, 1996.

[29] I. Cross. Is music the most important thing we ever did? music, development and evolution. In Suk Won Yi, editor, *Music Mind and Science*. Seoul National University Press, Seoul, 1999.

[30] Mihaly Csikszentmihalyi. *Flow: The Psychology of Optimal Experience*. Harper and Row, 1990.

[31] Mihaly Csikszentmihalyi. *The systems model of creativity: The collected works of Mihaly Csikszentmihalyi*. Springer, 2015.

[32] Sony CSL. Daddy's car: a song composed by artificial intelligence - in the style of the beatles. YouTube video, 2016.

[33] C. Darwin. *On the origin of species by means of natural selection, or The preservation of favoured races in the struggle for life*. D. Appleton and company, 1860.

[34] Richard Dawkins. *The Selfish Gene*. OUP, 1976.

[35] Ellen Dissanayake. *Art and Intimacy: How the Arts Began*. University of Washington Press, Seattle, USA, 2000.

[36] Emile Durkheim. *Selected Writings*. Cambridge University Press, New York, 1972.

[37] Jesse Engel, Cinjon Resnick, Adam Roberts, Sander Dieleman, Douglas Eck, Karen Simonyan, and Mohammad Norouzi. Neural audio synthesis of musical notes with wavenet autoencoders. *arXiv preprint arXiv:1704.01279*, 2017.

[38] Brian Eno. Generative music 1. Software art sleeve notes, 1996.

[39] Rebecca Fiebrink. Machine learning as meta-instrument: Human-machine partnerships shaping expressive instrumental creation. In *Musical Instruments in the 21st Century*, pages 137–151. Springer, 2017.

[40] Jamie Forth, Kat Agres, Matthew Purver, and Geraint A Wiggins. Entraining idyot: timing in the information dynamics of thinking. *Frontiers in Psychology*, 7, 2016.

[41] Bill Gaver, Tony Dunne, and Elena Pacenti. Design: cultural probes. *interactions*, 6(1):21–29, 1999.

[42] Alfred Gell. *Art and Agency: an Anthropological Theory*. Oxford: Clarendon Press, 1998.

[43] Pablo Gervás. A personal perspective into the future for computational creativity. In *Computational Creativity Research: Towards Creative Machines*, pages 393–406. Springer, 2015.

[44] Ignacio L Götz. On defining creativity. *The Journal of Aesthetics and Art Criticism*, 39(3):297–301, 1981.

[45] Edward H. Hagen and Gregory A. Bryant. Music and dance as a coalition signaling system. *Human Nature*, 14(1):21–51, 2003.

[46] Lejaren A Hiller Jr and Leonard M Isaacson. Musical composition with a high speed digital computer. In *Audio Engineering Society Convention 9*. Audio Engineering Society, 1957.

[47] Anna Katerina Jordanous. *Evaluating computational creativity: a standardised procedure for evaluating creative systems and its application*. PhD thesis, University of Sussex, 2013.

[48] Anna Kantosalo and Hannu Toivonen. Modes for creative human-computer collaboration: Alternating and task-divided co-creativity. In *Proceedings of the Seventh International Conference on Computational Creativity*, 2016.

[49] Anna Aurora Kantosalo, Jukka Mikael Toivanen, Hannu Tauno Tapani Toivonen, et al. Interaction evaluation for human-computer co-creativity. In *Proceedings of the Sixth International Conference on Computational Creativity*, 2015.

[50] Thomas S. Kuhn. *The Structure of Scientific Revolutions*. University of Chicago Press, 3 edition, 1996.

[51] Ray Kurzweil. *The age of spiritual machines: how we will live, work and think in the new age of intelligent machines*. Orion, 1999.

[52] Kevin N Laland, John Odling-Smee, and Marcus W Feldman. Niche construction, biological evolution, and cultural change. *Behavioral and brain sciences*, 23(01):131–146, 2000.

[53] Landr, 2020.

[54] Bruno Latour. *We have never been modern*. Harvard University Press, 2012.

[55] Thor Magnusson. Designing constraints: Composing and performing with digital musical systems. *Computer Music Journal*, 34(4):62–73, 2010.

[56] Lambros Malafouris. At the potter's wheel: An argument for material agency, May 2007.

[57] John Maynard Smith and Eörs Szathmáry. *The Major Transitions in Evolution*. Oxford University Press, New York, 1995.

[58] Geoffrey Miller. Evolution of human music through sexual selection. In Nils L. Wallin, Björn Merker, and Steven Brown, editors, *The Origins of Music*, pages 329–360. MIT Press, Cambridge, MA, USA, 2000.

[59] G Murphy. Typicality and the classical view of categories. *The Big Book of Concepts*, pages 11–41, 2004.

[60] Jakob Nielsen. 10 usability heuristics for user interface design. *Nielsen Norman Group*, 1(1), 1995.

[61] Donald Norman. *The Psychology of Everyday Things*. Basic Books, 1988.

[62] Adrian C North and David J Hargreaves. Music and adolescent identity. *Music education research*, 1(1):75–92, 1999.

[63] Francois Pachet. Beyond the cybernetic jam fantasy: The continuator. *IEEE Computer Graphics and Applications*, 24(1):31–35, 2004.

[64] Steven Pinker. *The language instinct: The new science of language and mind*, volume 7529. Penguin UK, 1995.

[65] Vilayanur S. Ramachandran. The artful brain. Talk given at the 2003 BBC Reith Lectures, available from http://www.bbc.co.uk/radio4/reith2003/lecture3.shtml, 2003.

[66] Mitchel Resnick, Brad Myers, Kumiyo Nakakoji, Ben Shneiderman, Randy Pausch, Ted Selker, and Mike Eisenberg. Design principles for tools to support creative thinking. 2005.

[67] Peter J. Richerson and Robert Boyd. *Not by Genes Alone: How Culture Transformed Human Evolution*. University of Chicago Press, Chicago, IL, US, 2005.

[68] Graeme Ritchie. Some empirical criteria for attributing creativity to a computer program. *Minds and Machines*, 17(1):67–99, 2007.

[69] R. Saunders. *Curious Design Agents and Artificial Creativity*. PhD thesis, Faculty of Architecture, The University of Sydney, 2001.

[70] Jürgen Schmidhuber. Formal theory of creativity, fun, and intrinsic motivation (1990–2010). *IEEE Transactions on Autonomous Mental Development*, 2(3):230–247, 2010.

[71] Jürgen Schmidhuber. Deep learning in neural networks: An overview. *Neural networks*, 61:85–117, 2015.

[72] Ben Shneiderman, Gerhard Fischer, Mary Czerwinski, Mitch Resnick, Brad Myers, Linda Candy, Ernest Edmonds, Mike Eisenberg, Elisa Giaccardi, Tom Hewett, et al. Creativity support tools: Report from a US national science foundation sponsored workshop. *International Journal of Human-Computer Interaction*, 20(2):61–77, 2006.

[73] Dan Stowell, Mark D Plumbley, and Nick Bryan-Kinns. Discourse analysis evaluation method for expressive musical interfaces. In *NIME*, pages 81–86, 2008.

[74] Bob L Sturm. The "horse" inside: Seeking causes behind the behaviors of music content analysis systems. *Computers in Entertainment (CIE)*, 14(2):3, 2016.

[75] Bob L Sturm, Oded Ben-Tal, Una Monaghan, Nick Collins, Dorien Herremans, Elaine Chew, Gaëtan Hadjeres, Emmanuel Deruty, and François Pachet. Machine learning research that matters for music creation: A case study. *Journal of New Music Research*, 48(1):36–55, 2019.

[76] Michael Tomasello, Malinda Carpenter, Josep Call, Tanya Behne, and Henrike Moll. Understanding and sharing intentions: the origins of cultural cognition. *Behavioral and Brain Sciences*, 28(5):675–691, 2005.

[77] Dan Ventura. The computational creativity complex. In *Computational Creativity Research: Towards Creative Machines*, pages 65–92. Springer, 2015.

[78] Dan Ventura. Mere generation: Essential barometer or dated concept. In *Proceedings of the Seventh International Conference on Computational Creativity*, pages 17–24. Sony CSL, Paris, 2016.