DESIGN CONSIDERATIONS FOR REAL-TIME COLLABORATION WITH CREATIVE ARTIFICIAL INTELLIGENCE

A PREPRINT

Jon McCormack*
SensiLab
Monash University
Caulfield East, Victoria, Australia
Jon.McCormack@monash.edu

Patrick Hutchings
SensiLab
Monash University
Caulfield East, Victoria, Australia
Patrick.Hutchings@monash.edu

Toby Gifford
SensiLab
Monash University
Caulfield East, Victoria, Australia
Toby.Gifford@monash.edu

Matthew Yee-King
Goldsmiths
University of London
London, UK
m.yee-king@gold.ac.uk

Maria Teresa Llano
Goldsmiths
University of London
London, UK
m.llano@gold.ac.uk

Mark d'Inverno
Goldsmiths
University of London
London, UK
dinverno@gold.ac.uk

February 8, 2020

ABSTRACT

Machines incorporating techniques from artificial intelligence and machine learning can work with human users on a moment-to-moment, real-time basis to generate creative outcomes, performances and artefacts. We define such systems collaborative, creative AI systems and in this paper, consider the theoretical and practical considerations needed for their design so as to support improvisation, performance and co-creation through real-time, sustained, moment-to-moment interaction. We begin by providing an overview of creative AI systems, examining strengths, opportunities and criticisms in order to draw out the key considerations when designing AI for human creative collaboration. We argue that the artistic goals and creative process should be first and foremost in any design. We then draw from a range of research that looks at human collaboration and teamwork, to examine features that support trust, cooperation, shared awareness and a shared information space. We highlight the importance of understanding the scope and perception of two-way communication between human and machine agents in order to support reflection on conflict, error, evaluation and flow. We conclude with a summary of the range of design challenges for building such systems in provoking, challenging and enhancing human creative activity through their creative agency.

Keywords Co-creation · creative activity · improvisation · performance · design · evaluation

1 Introduction

In recent years Artificial Intelligence (AI) has found an increasing role in the creative arts, indicating a growing interest amongst researchers and artists to collectively explore the possibilities. The field of *Creative AI* broadly looks at how AI techniques can be applied to the design and generation of creative artefacts, or to support human creators in their creative practices.

This paper considers a particular subset of AI technologies for music that enable real-time interactive improvisation, co-creation and performance to take place live between people and machines. We address a range of theoretical and methodological issues for designing and developing these systems. By doing so we hope to support the development of a stronger design framework for this exciting sub-field of AI for creative applications.

*Corresponding author

Through our interactions with these systems, our perception of them shifts from being tools or instruments that passively support human creativity, to a new kind of active creative partner. In this mode, the machine initiates new creative ideas, supports the human performer's creative practice in meaningful ways, and develops and adapts to the individual mode and style of artistic creation "in the moment" with a specific human partner. While mechanical and computational tools have, in the past, supported these concepts individually, we argue that it is the combination of these modes of engagement that leads to the perception of the machine as an "intelligent" creative partner.

Our interest in this class of real-time AI systems for creative collaboration – whose design requires combining artistic challenges and scientific goals – has several motivations:

- Performance and improvisation are amongst the most challenging creative activities undertaken by humans. To do them successfully requires a great deal of proficiency and virtuosity which typically takes many years of practice and experience before one can claim anything close to mastery [1, 2]. For machines to operate successfully in this situation provides us with a clear challenge, one that is quite unlike that of automatically generating finished artworks, mimicry of an existing style, or classifying creative outputs.
- For the first time in history, technology can provide artists with an opportunity to improvise, perform and co-create with intelligent machines. Co-creative machines suggest a new kind of agency: one that enables interactions that are fundamentally new and different from our previous creative interactions with human-made tools. This opens up exciting opportunities for human-machine cooperation that have never before been possible.
- The successful design of systems that have to create in real-time requires us to consider specific aspects of the human creative process in new ways. These include issues of teamwork, trust, cooperation, shared knowledge, social accountability, shifting goals and evaluation "in the moment".

These challenges provide us with a reality check on the limitations and affordances we require from AI to support and understand human creative processes. Moreover, they operate in contexts which are highly familiar to audiences, artists and performers alike.

1.1 A Practice-driven Approach to Creative AI

Artists have long used technologies to help them create, whether that is the pen, the piano or the computer. While technology often plays an important role, in our view the artistic work's meaning and relevance is foremost, regardless of the technology, methodology or tool being used. This means that it is important to begin with the artist's creative goals and understand how technology can assist them rather than the other way around. For creative works to have sustained appeal they should be driven primarily by artistic, not technological goals: what can be termed a *practice-driven* approach.

Using the latest AI techniques and demonstrating their effectiveness by generating outputs is insufficient without understanding the creative intention of human artists and audience expectations built around them [3, 4, 5, 6]. So we consider creativity not only determined by examining the final output, but through the process and experience of making it as well.

Despite AI's enormous creative potential, many remain sceptical or wary of it's ability to play any useful role in human creative practice or shed new light on our creativity beyond surface mimicry [3]. Issues for current creative AI research include:

- Over emphasising the final product rather than the underlying creative process;
- Developing techniques for single, specific individuals, performances or outcomes, that are not designed to elicit sustained interest or longer-term creative development and do not shed light on the design of collaborative AI systems in general;
- Using systems that cannot articulate or explain the decisions they have made, making understanding by both performers and audiences opaque and difficult.

To address these issues, we need a design process specifically oriented around collaboration and teamwork between human and machine performers. Real-time co-creative systems provide the opportunity for artists to develop collaborations with a creative agency rather than just work with a tool. Such collaborations require sustained, multiple encounters where each participant – human and machine – learns about the other. These interactions aim to build trust and familiarity through each exchange, balanced with a mutual openness to go into unexplored territory.

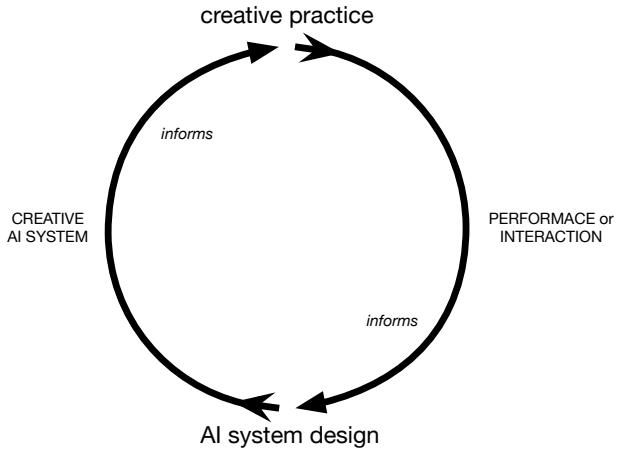

Figure 1: Creative practice and AI system design work best when they mutually inform each other in human-machine collaborations.

1.2 Artistic and Technological Co-development

Artistic practice is not static, it develops and changes over time, driven by culture, economics, personal development, experience and technology [7, 8, 9]. We think the most successful outcomes are those where the boundaries of creative practice and AI system design are simultaneously expanded by interplay with each other (Fig. 1). In this view, AI has multiple roles to play, one critical role is as a "cultural influencer" where, for example, new creative practices that are unique to human-AI collaborations are adopted across different creative domains (a type of *transformational creativity* in Boden's [10] terminology). These practices further inform the design of new AI systems, and the process repeats, leading to a symbiotic co-development between artistic and technological imperatives.

Beyond cultural change, there is evidence to suggest the field can have a wider societal relevance. In contemporary society, "creativity" is a highly sought-after commodity for both business and education. Regular music learning, for example, is associated with increases in IQ and general academic achievement [11, 12] and creative thinking abilities have been posited as the key to overcoming decline and stagnation for new economies [13].

In popular media, there is currently a huge interest in AI and what it might be able to achieve in the near future. But there is also anxiety about AI's potential negative consequences, such as job redundancies, biased decision making and loss of privacy. Opportunities for the public to better understand the reality of current AI research can be assisted through familiar scenarios such as music performance and improvisation. These use-cases demonstrate people working productively with AI, rather than being subsumed, beaten or replaced by it. They emphasise the positive value of AI and how it can contribute to a richer culture.

1.3 Overview of the Paper

We now have the opportunity to improvise, perform and co-create with intelligent machines. We think that these kinds of creative interactions are fundamentally new due to the increased scope for *agency* and *autonomy* that computing brings over previous human-made tools [14]. We will examine the relationships between improvisation, interaction and co-creation with an AI system in detail, specifically targeting AI techniques that support performing and co-creating with a creative human musical partner.

Section 2 provides an overview of Creative AI systems, illustrating the range of applications actively researched. This section also looks at some of the criticisms of Creative AI from creative arts and critical perspectives. Following that, Section 3 sets out in more detail what we believe are the main theoretical and technological design issues that need to be addressed when designing collaborative AI Systems. Drawing from research across psychology, musicology, improvisation theory and social informatics, we explore what can be learnt from these disciplines when designing AI systems that support real-time human collaboration with shared creative goals.

In the final section we discuss the important considerations for a methodological framework, assisting the field to develop in more coherent and informed ways.

2 Background and Related Work

The application of AI methods for real-time musical performance and improvisation has become prominent of late, due to increased accessibility of open source, re-usable software components for analysis and classification using techniques such as Deep Learning, combined with the increase in readily available compute resources and large datasets. However, real-time music performance systems have existed for nearly half a century [15]. Gifford et al. undertook a survey of computational systems for music improvisation, and developed a taxonomy through a detailed examination of 23 indicative systems, covering all major approaches [16]. Their key findings included the idea that system complexity had little influence on the perceived creative agency of the artificial improviser and the conceptualisation of the system as a creative partner dates back more than 30 years.

2.1 Artificial Creativity

Many different AI techniques have been applied to tasks typically associated with human creativity, including the production of visual and moving image art [17], literature and music [18], but also more broadly in areas such as games [19] and game shows [20].

The development of artificial creative systems has generally been a bespoke and multifaceted effort driven by several competing objectives. These range from understanding human creativity, to assisting human creators through the practical generation of specific artefacts, to building systems that are deemed – to varying degrees – as autonomously creative. A recurring challenge for AI is "to beat the best human" at some specific task or problem. But this approach seems counter to the goal of expanding and supporting all human creativity.

An alternative is to develop new models of interaction and co-creation that are designed to nurture and enhance the user's creativity and creative practice. In this mode of *AI-as-collaborator*, human-machine collaborations have shown to foster human creativity in specific contexts [21, 22]. One approach is to parameterise systems for the generation and curation of output. In music composition collaborators specify parameters such as the types of melodies and themes for generated compositions to be based on, time and key signatures to constrain the number of melodies, and so on. [23, 24]. Higher levels of autonomy have also been seen in systems that make decisions based on metadata-level criteria. For example, the goal-awareness approach of [25] models its collaborators by measuring their musical potential in order to make decisions about how to carry on with the collaboration.

2.2 Beyond Mimicry: Collaborative Co-creation

A recurring question for many creative AI systems, trained on an existing artistic corpus or style, is whether they have simply learned how to imitate an artist's style but cannot go beyond it. An alternate approach is to assist with the generation of new artefacts rather than mimicry of existing artistic styles [26]. Researchers can imbue their systems the basic aspects of music theory – such as the way some combination of sounds shared a particular meaning and trigger particular emotions – in order to provide a general tool that musicians can use to create soundtracks, rather than relying only on big data analysis of famous composers' corpus or using machine learning methods to develop statistical models of musical genres [27].

These kind of systems have enabled successful creative collaborations with artists.[2] *Flow Machines* [28] the technology behind *Daddy's Car* ("the first structured AI pop song"), is a prominent example of AI for musical co-creation. In such systems musical styles – including melodic phrases and harmonisation and progression patterns – are represented as computational objects, allowing users to explore, manipulate and change styles in order to create their own.

Research has emphasised the value in developing systems that focus on the creative process rather than just the artefact output [29]. For example by giving more autonomy to the creative process by allowing systems to "inspire themselves" as part of the artefact generation process [30], or enabling them to write their own code [31]. The underlying rationale is that by allowing greater autonomy into the decision making process, we (and therefore our AI systems) can better understand what aspects of the creative process lead to successful outcomes. Some of these interactions are criticised because the software "merely follows a random process" to generate suggestions [21], and it is actually the artist's decisions that shape the space from which the software generates alternatives. Nonetheless, researchers claim these approaches increase the distance between developer intentions and the system's creative process, even possibly uncovering novel and interesting ways of artefact generation (in spite of the resulting artefact being possibly of poor quality – quite common also in human creative practices).

[2] `https://www.theverge.com/2017/8/27/16197196/taryn-southern-album-artificial-intelligence-interview`

2.3 Evaluation

> Creative activity is our great need, but criticism, self-criticism is the road to its release.
>
> — John Dewey [32].

To deem a system "creative" and for it to improve and grow creatively, evaluation is necessary [33]. This may consist of machine reflection and self-criticism, or may arise through mutual interplay and feedback of a human-machine partnership where the human is performing real-time aesthetic evaluation.

Agres and colleagues [34] classified creative systems according to the degree to which they can reflect about their own output, suggesting three main categories: (i) purely generative; (ii) with internal or external feedback; and (iii) with capabilities of reflection and self-reflection.

External feedback through human-machine interaction can be considered as negotiated evaluation of the creative partnership. Additionally, the primary mode of communication may in fact be via the creative medium itself. In human-human music improvisation it is through the music that communication and negotiation primarily takes place. Whilst a plethora of extra-musical communication channels are involved, such as physical gestures, eye contact, and even verbal cues, these are often seen as secondary across jazz [35], free [36], and electroacoustic [37] improvisation genres. Similarly, interactive music systems such as *Cypher* [38], *OSCAR* [39], *Voyager* [40] and *CIM* [41] privilege this mode of "performance-as-interface" [42] whether or not some additional parametric controls are exposed. Thus human evaluation in the human-machine creative partnership can enter through the creative improvisation itself. The challenge is for the machine to successfully interpret and to act on it so as to improve.

User experience can form part of the evaluation of creative systems, and recent studies have supported a co-creational approach by placing emphasis on usability evaluation, specifically geared towards a subjective and experiential analysis rather than to task-based usability found in traditional HCI research [43]. Accordingly it is important to consider how *human* improvisations are evaluated. In iterative models of the creative process, evaluation plays a key role in the iterative cycle and the creator is often in a constant process of evaluation [44, 45, 33].

One important methodological consideration that has received much support from researchers is the use of subjective aesthetic assessment by "domain insiders" as an instrument for evaluating improvisation [46, 47, 48]. This parallels similar arguments in creativity assessment theory [49, 50].

In designing AI collaborators, it may be that the designer themselves plays the role of expert judge. For example, in the context of designing Digital Musical Instruments (DMIs), Jordá and Mealla comment:

> [Just as] much research in HCI culminates in lists of guidelines and/or principles for design (and/or evaluation of design) based on research or practical experience relating to how people learn and work, it comes as no surprise that the first tentative NIME design frameworks have been mostly proposed by experienced digital luthiers [51].

As all creative domains have some culture of critique, and corresponding critical values, embedding criticism in AI design seems an obvious fit. However, making those values explicit remains challenging. As Eisenberg and Thompson note, "the evaluation of improvised music is particularly mysterious" [50]. Consequently implicit approaches leveraging the subjective real-time evaluation of a "human-in-the-loop" are fundamental to good system design.

2.4 AI-as-Collaborator

There has been a significant amount of research into designing competent improvisational companions. The goal of these systems is to contribute creatively during performances in order to stimulate the human collaborator(s) into new expressions or creative territory. Approaches include the use of a predefined palette of subroutines [52, 53, 54], grammars [55, 56], or a combination of both [57]. While these approaches require a human expert to design the subroutines or grammar rules, other approaches employ machine learning of performance patterns, which are then used during real-time performance to generate novel compositions [58, 59, 60, 61]. Collectively, these systems tend towards being bespoke to their designer(s), making general or sustained creative development difficult (c.f. §1.1). Our own projects have incorporated AI techniques for real-time co-creational applications, including *Eden* [62], *Reflexive Looper* [63], and *Controlling Interactive Music (CIM)* [42].

2.5 Summary: Towards a Truly Collaborative AI

The notion of AI-as-collaborator is being increasingly studied within the field of creative AI, becoming (in our view) a fruitful and productive approach as evidenced by the creative sophistication and audience acceptance of results [64].

However, many efforts in creative AI are driven by the goal of attaining expert-level human performance rather than by understanding how a system can support and enhance human creative activity collaboratively. While this has driven important technological advances, in order to advance the concept of AI-as-collaborator we need to follow a practice-driven approach (§1.1), which furthers human creative practice as its first and foremost design goal. Then only by challenging, provoking, stimulating and pushing the process and experience of human creative activity are new artistic achievements possible.

Steps are now being taken towards this goal: researchers have begun moving away from approaches that simply mimic or parody existing artistic canons, focusing instead on the creative process rather than the output and favouring evaluation through self-reflection and ecological experience over traditional task-based measurements.

Improvisation, performance and co-creativity provide ideal activities in which to explore the role of machines as creative partners, i.e. creative collaborators that are engaged with an artistic goal but that have the freedom to explore different possibilities to achieve it.

In the next section we explore some of the considerations that can assist in building models that foster an ongoing, sustained dialogue for real-time human-machine collaborations. It seems logical that human-machine partnerships can be better understood through the lens of how humans collaborate with one another. To this end, we study collaborations from various perspectives in the human context, mainly surrounding the role of team work (or lack of thereof) within these interactions, and propose how these different perspectives can be applied to the design of AI collaborators.

3 Design Considerations for Collaborative AI systems

Here we present a set of design considerations for creating collaborative real-time AI systems. We take the position that collaboration between AIs and people shares characteristics with collaboration between people, and therefore that the nature and mechanisms of successful human teams can inform the design of AI collaborators.

We begin by considering the characteristics of a dysfunctional team, based on a widely used model from management studies put forward by Lencioni. We contextualise this model by imagining a perfectly dysfunctional band of musicians, then a perfectly dysfunctional AI collaborator.

Based on this knowledge of how teams can fail, we develop a set of considerations for people interested in designing AI systems that can operate successfully as team members in creative collaborations.

3.1 How do dysfunctional teams behave?

Lencioni, a theorist in management studies, identified five areas of dysfunction in teams: absence of trust, fear of conflict, lack of commitment, avoidance of accountability and inattention to results [65].

Imagine a band of musicians who exhibit all these problems. According to Lencioni, lack of trust leads to undesirable behaviours such as concealment of weaknesses. The drummer does not admit that they cannot play in 7/8 time, and refuses to explain why they come in at the wrong time with the wrong beat. Fear of conflict leads to stagnant music because nobody is prepared to criticise anyone else's playing, and therefore to encourage improvement. Lack of commitment leads to nobody taking the instrumental solo as they want to get through the performance with minimal risk. Avoiding accountability means nobody wants to admit why the music is stagnant, or why the drummer came in at the wrong time or why nobody took the solo. Finally, inattention to results leads to the band walking off stage at the end and none of them commenting on the performance and why it was so terrible, or why the audience left early. Even the success of the guitarist, who is very goal driven and who set fire to his guitar and destroyed his amplifier very effectively, is not commended by his team-mates in the van on the way home.

A dysfunctional band of musicians is one thing, but what would be the experience of interacting with a Lencioni dysfunctional AI collaborator? We shall consider a scenario in which a person is co-composing a piece of music with an AI. The AI should be able to provide feedback on the musical score and to suggest changes. What kind of suggestions might an untrustworthy AI make? It might heap positive praise on everything the human writes, or it might criticise without justification. If the AI fears or is programmed to avoid conflict, it avoids any critique or suggestion at all if it challenges the human musician. Since the AI does not contribute any real critique to the composition, it has no "skin in the game", and therefore bears no responsibility or accountability for the results. Finally, the AI attends to all results with the same, positive response, so it does not recognise the significance of different results, be they good or bad music.

So what makes for a good team and how should we design collaborative AI systems, so they are good team members? We have explored the characteristics we should avoid, but let us now consider some other research into teamwork that explains mechanisms we can use to help us avoid these problems.

3.2 Designing for trust

A strong and productive human creative partnership is fundamentally founded on trust between the participants: performers trust their collaborators to share, contribute and participate towards a common creative goal. Yet this goal isn't always formalised or articulated explicitly beforehand, rather it is often negotiated during performance. This makes it very difficult to present it to an AI for consideration. An improvisational collaboration takes place "in the moment" of performance – in real-time with little or no time for conscious reflection or planning – meaning that it becomes very easy to disrupt if performers are not all invested in an overall creative outcome. Hence, trust in an exemplary performance plays over many different relationships and levels: trust in each participant's *intention* and creative virtuosity [66]; trust that each articulation has a purpose in the overall performance; trust in the expectations of a response put "out there" by another, trust in the understanding of the current creative direction of the work as it progresses, trust that allows participants "to risk everything in the moment of performance" [67].

A lack of trust in an improvisational partner can foster fears of failure in performance, amplifying worries about one's own ability to perform competently with others or undermining the willingness to act freely. On the other hand, performers must take risks that embolden them into new creative ideas and territories rather than remain in stasis or isolation during a performance or improvisation. Many see this negotiation between trust and risk within improvisation and collaborative performance as social interaction and accountability [67], which raises interesting questions on how to ensure trust between humans and creative machines. Certainly, trust in computing and automated systems is well studied, where similar issues of social interaction and accountability are acknowledged as fundamentally important [68]. Typically however, social interactions with machines are asymmetrical, with an unbalanced awareness of each other's behaviour and intentions [69], making building the necessary trust problematic. As more complex AI systems become part of those interactions, this asymmetry may increase.[3]

Trust can be attributed through direct observation of three layers of abstraction: *performance* (behaviour), *process* (underlying mechanisms) and *purpose* (system intent). Building trust with a non-human agent also requires *calibration* between a person's expectations of the agent and the agent's capabilities [70, 71, 68]. Exposing these three layers in any collaborative AI system can serve as a useful design goal, with the constraint of designing for an appropriate level of trust rather than trying to maximise it. An "appropriate" level balances trust in the system with the necessary risk to push the interaction creatively.

Another complex factor in building trust is the use of anthropomorphisms. Over use of anthropomorphic skueomorphisms, such as human-like voices, facial expressions, bodies, etc. leads to false assumptions about underlying mechanisms and capabilities of the system, because we expect it to be more human-like than it really is. This parallels the "uncanny valley" effect well known in animation, though originally discussed in the context of humanoid robotics [72].

While impressions of creative trust in a performative or improvisational context can be evaluated effectively, formalising them from an AI perspective is problematic. Hence we see a measure of trust as a possible evaluative outcome of a human-machine co-creation, rather than something that is directly built into a design in an engineering sense. Some possible mechanisms for building trust in human-machine partnerships are outlined in the sections that follow.

3.3 Designing for team cognition

We can think of machine-human improvisation or co-creation as teamwork, which is a well-studied area in psychology. What, according to psychologists, are the factors that affect the success of teams and how might we consider them in the design process for collaborative AIs? In a meta-analysis of the cognitive underpinnings of effective teamwork, DeChurch et al. highlight the importance of *team cognition* [73]. They identify two critical elements of team cognition: team mental models and transactive memory.

Team mental models are defined as "organized mental representations of the key elements within a team's relevant environment that are shared across team members" [74]. Essential elements here are the roles and capabilities of the team members and the expected procedure for the work.

Wegner defines transactive memory as a "cognitively independent system for encoding, storing and retrieving information that combines the knowledge possessed by each individual with a shared awareness of who knows what." [75].

[3] This is particularly the case for methods such as deep learning, where the reverse engineering of a complex neural networks operation becomes intractable. While we face a similar problem with the human brain at a neurological level, we are at least able to self-articulate at the higher level of intentions, beliefs and goals to others.

Trust will emerge only when team members are confident that there are shared models and shared information, so they need to be there, and the team need to know they are there.

The ideas of shared dynamic models and a shared space for information storage (which is how one might paraphrase team mental models and transactive memory) are also apparent in the literature on music improvisation. In a review paper considering the literature relating specifically to free improvisation, Ng identifies the themes of situated, collaborative knowledge construction, and the real-time emergence of means of working towards shared goals [76]. It seems that musical interaction theorists take the ability to share information and to be aware of others' goals and capabilities for granted in musical interaction, suggesting an implicit acceptance of shared mental models and shared information. The theory of social constructivism seems to neatly encompass shared models, shared information and the emergence of a meaningful, trusted collaboration [77].

So how do we unlock this implicit information and expose it to team members in a human-AI collaboration? Lewis, who developed the Voyager computer music improviser considered how computer systems might communicate their internal state to human band members during an improvisation. He stated that "the nature of the internal representation used by a system will be audible to the trained improviser based on the system's performance and the improviser's experiences with it" [40]. Therefore the shared awareness required for successful teamwork does not necessarily need to be explicitly expressed, e.g. through visualisation or numerical display, as it will be apparent in the output the system is designed to generate (music in this case).

In summary, we are interested in systems which work alongside people to enhance, provoke and challenge their creativity. In real-time, interactive contexts, trust is a critical element of the interaction, and it is a complex phenomenon to describe and model. We do know that trust emerges when the collaborators are aware that they have shared goals and direction.

Psychologists interested in teamwork have described the mechanisms and content of these shared goals and how they are communicated, in the form of shared dynamic models of the team members' roles and capabilities, shared awareness of the work flow they need to engage in to achieve the task at hand, and some sort of shared information space.

3.4 Designing for feedback and discussion during the creative process

As discussed in §1, a common issue with computational creative systems is that their sole focus is on the end product. This reduces the interaction between a person and the AI to a request followed by an acceptance or rejection, rather like ordering something from an on-line shop, or Thorndike's Stimulus-Response model of learning.

Producing an artefact is only one of many possible creative and artistic activities. People create various versions of their work, discuss ideas, intentions, influences about their work, compare them to previous works, and so on. These types of communication have been identified as an intrinsic component of creativity, "attempting to communicate a creative work often feeds back to fundamentally transform the creative work itself." [33]. Enabling systems with communication and dialogue capabilities is thus increasingly seen as paramount. This dialogue is also an enabler of trust, as opinions are presented, justified and interrogable. Lencioni tells us that the opportunity to confront is critical to excellent teamwork, as is accountability, both of which are enabled by dialogical systems.

We can imagine an AI system that holds a useful mental model of its human collaborator and which can reason about that model and express that reasoning. Such a system would need to understand the domain of work, and the conventions governing how one operates in that domain. For example, imagine a musical band-leader demonstrating on the piano how the pianist should emphasise a particular phrase – that is quite normal. Now imagine an AI painter-collaborator that paints all over the section of the painting they think should be emphasised.

Effective communication would enable systems to interact with human collaborators at new levels. A system that can communicate its mental models and the mental models it holds of its collaborators would engender increased trust from these collaborators (See § 3.3). As described in [23], "collaboration embraces flaws", and two-way communication would also provide systems with the ability to reflect about conflicts, mistakes and past versions. In this kind of model, the systems would become active participants as opposed to only supporting their human collaborators. Finally, having the ability to explain their contributions would also mean giving collaborative AI systems the ability to argue for them and achieve a more balanced partnership.

3.5 Designing for Agency, Autonomy and Reflection

Like intelligence, creativity has historically been difficult to define, and indeed, has had different definitions at different points across the history of the term's use [4]. Considerations of agency, autonomy and reflection provide the designer with more straightforward ways to characterise the nature of the interaction between human and machine.

Recent work unpicks a distinction between the autonomous heroic lone agent that produces content without interaction with human agent, and the collaborative creative computing agent that interacts with a human user to create new kinds of behaviours and performances [5]. It outlines the differences between the two approaches from the of the human artist/performer, the aims and motivations of the human designer/software-engineer and the experiences of an audience. This approach draws from concepts of agency and autonomy [78] where non-autonomous agents serve the (in our case artistic motivations) of another, and autonomous agents have their own motivations and so can generate their own artistic goals. Currently, questions about the level of an AI's autonomy is primarily used in the characterisation of systems rather than in directing the creative process and generating artistic or performative "goals", what we have termed "creative agency" [79]. Shifting the emphasis to both developing and supporting creative goals seems a more beneficial approach.

3.5.1 Reflection

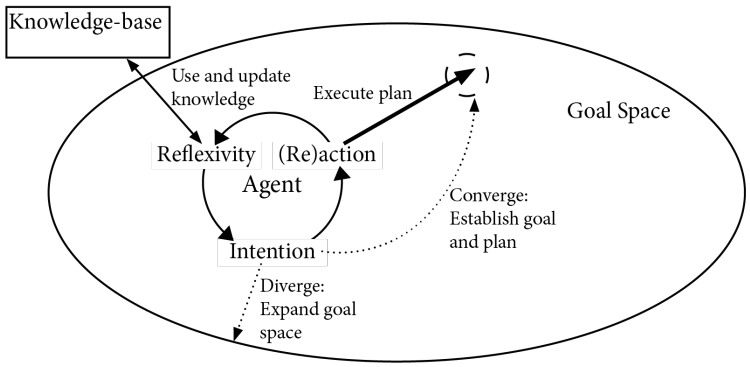

Figure 2: The internal processes that drive agency and autonomy in collaborative agents modify the "goal space" which is expanded through divergent thought and reflected on through action and reference to knowledge. These reflections guide intentions to converge on specific goals, establish plans and act on them.

We consider reflection the ability of an agent to look back upon the process and results of its collaborations. Currently, designs rarely consider the opportunity to grow and adapt through evaluation and reflection of past successes or failures. The majority of creative systems are switched on to create something and then stopped [80]. We can connect this design flaw to one of the characteristics of dysfunctional teams – inattention to results – emphasising the importance of considering reflection in the design of collaborative AI systems.

One way to focus on process and encourage growth is to reconceptualise creative agent systems from a two-step, action-reaction model to agents as systems of continuous, cyclical co-ordination (Fig. 2). The term "co-agency" has been used to describe the bi-directional relationship of agents with *"cyclical coordination between intentionality, (re)action and reflexivity"* [81]. This model highlights a limitation of current creative systems, in that they are typically excluded from establishing intention.

3.5.2 Goals and Intention

Both agency and autonomy pertain to goals, either in how they are formed, or how they are attempted and solved. Goals can only be formed with the knowledge the system has access to and intention can only be established with the tools to form plans to reach these goals and commitment to carry them out [82]. Autonomy to develop new goals is bounded by knowledge. As such, exploring knowledge from different domains and translating it to a desired context is a fertile area for expanding the goals of creative systems and giving them the tools to act with intention [83].

Interesting translations of knowledge are typically discovered by people and used to establish the knowledge base and tool-set of their systems. DeepDream [84] and neural style transfer [85] techniques utilise object detection models to generate interesting visual artefacts. AI systems with knowledge outside of the domain of application can help to expand the search for novel ideas and then narrow down those ideas by finding patterns and algorithms that have been useful or interesting in other domains.

Creating goals, and establishing intention to act on them can be analysed through the concept of "Convergent and Divergent thinking", where new ideas and possibilities are explored and then the best ideas are selected, focused on and developed. This concept has been used by psychologists to explain creative mental processes in individuals and groups [86, 87].

While divergent thinking is most commonly associated with creativity, with studies showing creative professionals have increased mental activity on tasks associated with divergent thinking [88], developing creative thoughts into tangible artefacts requires evaluation of novelty and realisation, making convergent thinking also essential in creative practice [89]. Balancing the flow of converging and diverging states, building goals and intention to reach them, is a skill that creative people excel at, making it an important process to consider in the design of creative AI systems.

4 Conclusion

There is little doubt that AI and machine learning are changing the way that improvisation and performance are conceptualised and realised. We have undertaken an analysis of the role of AI systems that interact with humans in real-time, on a moment-to-moment basis over a sustained period to produce creative content and encourage human creative development. We have synthesised a set of design considerations for researchers wishing to create AI systems that can collaborate with people. The key elements discussed are trust, team cognition, feedback and autonomous agency with reflection.

We are interested in these systems because they focus our attention on specific design challenges for the AI researcher *and* the creative practitioner: how and why would we interact with a machine? What is the payoff for the engineer, the artist co-creating with the system, and the audience? What can be generated that could not be generated without the active participation and creative agency of the machine? To what extent – if at all – does the audience need to know about the agency or design of the system in order to appreciate the content?

Throughout our discussions in this paper we have tried to show that it is important to understand the experience of the human co-creator and their relationship to the artificial creative partner, so as to design systems that have genuine creative agency (unlike the piano or pen) that can be sustained beyond a single performance or fixed interaction.

The considerations presented in this paper provide practitioners with considerations for analysing, describing and creating systems with individual merit, supporting a wide breadth of scientific and artistic aims whilst contributing to the field more generally. However, we recognise that this field is in its infancy. Many systems have been designed for a single performance or performer, rather than conceptualised as a creative collaborator that can build a sustained and maturing creative relationship like those which exist between successful human collaborators.

We have offered a set of design considerations, based around successful human collaborative teams, as simple but effective criteria from which to consider the design of collaborative AI systems. Our vision is for an expanded notion of creativity: one that supports new possibilities for the human creative artist through a productive partnership with AI, rather than trivialising, superseding or replacing human creativity. We look forward to seeing this vision unfold over the coming decades.

5 Acknowledgements

This research was support by Australian Research Council grants DP160100166 and FT170100033. Many of us wish to acknowledge our gratitude to Arthur Still for introducing us to the work of John Dewey.

References

[1] K.A. Ericsson, R. Krampe, and C. Tesch-Römer. The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100:363–406, 1993.

[2] M. Gladwell. *Outliers, the Story of Success*. Allen Lane, 2008.

[3] Anthony O'Hear. Art and technology: An old tension. *Royal Institute of Philosophy Supplement*, 38:143–158, 1995.

[4] Arthur Still and Mark d'Inverno. A history of creativity for future AI research. In F. Pachet, A. Cardoso, V. Corruble, and F. Ghedini, editors, *Title: Proceedings of the 7th Computational Creativity Conference (ICCC 2016)*. Universite Pierre et Marie Curie, 2016.

[5] Mark d'Inverno and Jon McCormack. Heroic versus collaborative AI for the arts. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015)*, pages 2438–2444. AAAI Press, 2015.

[6] Matthew Yee-King and Mark d'Inverno. Experience driven design of creative systems. In F. Pachet, A. Cardoso, V. Corruble, and F. Ghedini, editors, *Title: Proceedings of the 7th Computational Creativity Conference (ICCC 2016)*. Universite Pierre et Marie Curie, 2016.

[7] E. H. Gombrich. *The story of art*. Prentice-Hall, Englewood Cliffs, N.J., 16th, rev., expanded and redesigned edition, 1995.

[8] Cynthia A. Freeland. *But is it art?: an introduction to art theory*. Oxford University Press, Oxford, 2001.

[9] Sarah Thornton. *Seven Days in the Art World*. Granta, London, UK, 2009.

[10] Margaret A. Boden. *The creative mind: myths & mechanisms*. Basic Books, New York, N.Y., 1991.

[11] E Glenn Schellenberg. Music lessons enhance iq. *Psychological science*, 15(8):511–514, 2004.

[12] E. Glenn Schellenberg and Michael W. Weiss. Music and cognitive abilities. In *The Psychology of Music*, chapter 12, pages 499–550. Elsevier, Inc., 2013.

[13] Daniel H Pink. *A whole new mind: Why right-brainers will rule the future*. Penguin, 2006.

[14] Oliver Bown and Jon McCormack. Taming nature: tapping the creative potential of ecosystem models in the arts. *Digital Creativity*, 21(4):215–231, 2010.

[15] Arne Eigenfeldt. Real-time composition or computer improvisation? a composer's search for intelligent tools in interactive computer music. *Proceedings of the Electronic Music Studies*, 2007, 2007.

[16] Toby Gifford, Shelly Knotts, Jon McCormack, Stefano Kalonaris, Matthew Yee-King, and Mark d'Inverno. Computational systems for music improvisation. *Digital Creativity*, 29(1):19–36, 2018.

[17] L. Avila and M. Bailey. Art in the digital age. *IEEE Computer Graphics and Applications*, 36(4):6–7, 2016.

[18] David Cope. Experiments in musical intelligence (emi): Non-linear linguistic-based composition. *Interface*, 18(1-2):117–139, 1989.

[19] J. X. Chen. The evolution of computing: Alphago. *Computing in Science Engineering*, 18(4):4–7, 2016.

[20] Jo Best. IBM Watson the inside story of how the jeopardy winning supercomputer was born and what it wants to do next, September 9 2013.

[21] Antonios Liapis et al. Can computers foster human users' creativity? theory and praxis of mixed-initiative co-creativity. *Digital Culture & Education (DCE)*, 8(2):136–152, 2016.

[22] Jon McCormack, Toby Gifford, Patrick Hutchings, Maria Teresa Llano Rodriguez, Matthew Yee-King, and Mark d'Inverno. In a silent way: Communication between ai and improvising musicians beyond sound. In Anind Dey and Shengdong Zhao, editors, *Proceeding of ACM SIGCHI 2019*, page Paper No. 38, New York, NY, May 2019 2019. ACM SIGCHI, ACM.

[23] Margareta Ackerman et al. Co-Creative Conceptual Art. In *Proceedings of the Ninth International Conference on Computational Creativity*, 2018.

[24] C. Cassion et al. With bolts of melody! songwriting with alysia & emily dickinson. In *Musical Meta-Creation*. 2017.

[25] Otto Hantula and Simo Linkol. Towards Goal-aware Collaboration in Artistic Agent Societies. In *Proceedings of the Ninth International Conference on Computational Creativity, ICCC 2018*, 2018.

[26] Ahmed M. Elgammal et al. Can: Creative adversarial networks, generating "art" by learning about styles and deviating from style norms. *CoRR*, abs/1706.07068, 2017.

[27] Patrick Hutchings. *Adaptive Music Scores for Interactive Media*. PhD thesis, Monash University, 2018.

[28] Fiammetta Ghedini et al. *Creating Music and Texts with Flow Machines*, pages 325–343. Springer Singapore, Singapore, 2016.

[29] Takashi Iba. An autopoietic systems theory for creativity. *Procedia Social and Behavioral Sciences*, 2:6610–6625, 2010.

[30] Simon Colton. *The Painting Fool: Stories from Building an Automated Painter*, pages 3–38. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.

[31] M. Cook et al. The ANGELINA videogame design system – part I. *IEEE Transactions on Computational Intelligence and AI in Games*, 9(2):192–203, 2017.

[32] John Dewey. *Construction and criticism*, volume 1. Columbia University Press, 1930.

[33] R Keith Sawyer. *Explaining creativity: The science of human innovation*. Oxford University Press, 2011.

[34] Kat Agres et al. Evaluation of musical creativity and musical metacreation systems. *Comput. Entertain.*, 14(3):3:1–3:33, December 2016.

[35] Garry L. Hagberg. Jazz improvisation and peak performance: Playing in the zone. In Tim Jordan, Brigid McClure, and Kath Woodward, editors, *Culture, Identity and Intense Performativity*, chapter 10. Taylor & Francis, 2017.

[36] Tomas E Nunn. *Wisdom of the impulse: On the nature of musical free improvisation*. Thomas E. Nunn, 1998.

[37] Doug Van Nort. Conducting the in-between: improvisation and intersubjective engagement in soundpainted electro-acoustic ensemble performance. *Digital Creativity*, 29(1):68–81, 2018.

[38] Robert Rowe. *Interactive music systems: machine listening and composing*. MIT press, 1992.

[39] Peter Beyls. Introducing oscar. In *Proceedings of the International Computer Music Conference*. ICMA, 1988.

[40] George E Lewis. Interacting with latter-day musical automata. *Contemporary Music Review*, 18(3):99–112, 1999.

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

[42] Andrew R. Brown. Creative improvisation with a reflexive musical bot. *Digital Creativity*, 29(1):5–18, 2018.

[43] Dom Brown, Chris Nash, and Tom Mitchell. A user experience review of music interaction evaluations. In *NIME 17*, pages 370–375, Copenhagen, 2017.

[44] Michael D Mumford, Wayne A Baughman, and Christopher E Sager. Picking the right material: Cognitive processing skills and their role in creative thought. 2003.

[45] J. P. Guilford. *The nature of human intelligence*. McGraw-Hill series in psychology. McGraw-Hill, New York, 1967.

[46] Dan Stowell, Andrew Robertson, Nick Bryan-Kinns, and Mark D Plumbley. Evaluation of live human–computer music-making: Quantitative and qualitative approaches. *International Journal of Human-Computer Studies*, 67(11):960–975, 2009.

[47] Sile O'Modhrain. A framework for the evaluation of digital musical instruments. *Computer Music Journal*, 35(1):28–42, 2011.

[48] Adam Linson, Chris Dobbyn, and Robin Laney. Critical issues in evaluating freely improvising interactive music systems. In *International Conference on Computational Creativity*, page 145, 2012.

[49] Teresa M Amabile. *Creativity in context*. Westview Press, Boulder, Colo., 1996.

[50] Jacob Eisenberg and William Forde Thompson. A matter of taste: Evaluating improvised music. *Creativity Research Journal*, 15(2-3):287–296, Jul 2003.

[51] Sergi Jordà and Sebastián Mealla. A methodological framework for teaching, evaluating and informing nime design with a focus on expressiveness and mapping. In *NIME*, volume 14, pages 233–238, 2014.

[52] George E. Lewis. Too many notes: Computers, complexity and culture in voyager. 10:33–39, 12 2000.

[53] Toby Gifford and Andrew Robert Brown. Beyond reflexivity: Mediating between imitative and intelligent action in an interactive music system. In *25th BCS Conference on Human-Computer Interaction*, 2011.

[54] Andrew R. Brown et al. Stimulating creative partnerships in human-agent musical interaction. *Comput. Entertain.*, 14(2):5:1–5:17, January 2017.

[55] Robert Keller and David R Morrison. A grammatical approach to automatic improvisation. 01 2007.

[56] Jon McCormack. *Complex Systems 96: from Local Interactions to Global Phenomena*, chapter Grammar-Based Music Composition, pages 321–336. ISO Press, Amsterdam, 1996.

[57] Kris Kitani and Hideki Koike. Improvgenerator: Online grammatical induction for on-the-fly improvisation accompaniment. pages 469–472, 2010.

[58] John A Biles. Genjam: A genetic algorithm for generating jazz solos. In *ICMC*, volume 94, pages 131–137, 1994.

[59] B. Thom. Learning models for interactive melodic improvisation. pages 190–193, 1999.

[60] Gérard Assayag and Marc Chemillier. Omax brothers : a dynamic topology of agents for improvization learning. In *Proceedings of the 1st ACM workshop on Audio and music computing multimedia*, 2006.

[61] G. Weinberg et al. Interactive jamming with shimon: A social robotic musician. In *2009 4th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 233–234, 2009.

[62] Jon McCormack. The evolution of sonic ecosystems. In Maciej Komosinski and Andrew Adamatzky, editors, *Artificial Life Models in Software*, pages 393–414. Springer, London, 2nd edition, 2009.

[63] François Pachet, Pierre Roy, Julian Moreira, and Mark d'Inverno. Reflexive loopers for solo musical improvisation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2205–2208. ACM, 2013.

[64] Gerfried Stocker. Aixmusic festival'19. Festival Program, Ars Electronica 2019, Festival for Art, Technology & Society, Linz, Austria, September 2019.

[65] Patrick Lencioni. *The five dysfunctions of a team*. John Wiley & Sons, 2006.

[66] François Pachet. Musical virtuosity and creativity. In *Computers and creativity*, pages 115–146. Springer, 2012.

[67] Ellen Waterman. Improvised trust: Opening statements. In Rebecca Caines and Ajay Heble, editors, *The Improvisation Studies Reader: Spontaneous Acts*, chapter 8, pages 59–62. Routledge, Oxon, UK, 2015.

[68] John D. Lee and Katrina A. See. Trust in automation: Designing for appropriate reliance. *Human Factors*, 46(1):50–80, 2004.

[69] M. Deutsch. The effect of motivational orientation upon trust and suspicion. *Human Relations*, 13:123–139, 1960.

[70] B. M. Muir. Trust between humans and machines, and the design of decision aides. *International Journal of Machine Studies*, 27:527–539, 1987.

[71] J. D. Lee and N. Moray. Trust, self-confidence, and opera- tors' adaptation to automation. *International Journal of Human-Computer Studies*, 40:153–184, 1994.

[72] M. Mori, K. F. MacDorman, and N. Kageki. The uncanny valley. *IEEE Robotics Automation Magazine*, 19(2):98–100, June 2012.

[73] Leslie A. DeChurch and Jessica R. Mesmer-Magnus. The cognitive underpinnings of effective teamwork: A meta-analysis. *Journal of Applied Psychology*, 95(1):32–53, 2010.

[74] S. Mohammed et al. Metaphor no more: A 15-year review of the team mental model construct. *Journal of Management*, 36(4):876–910, 2010.

[75] Daniel M. Wegner. Transactive Memory: A Contemporary Analysis of the Group Mind. *Theories of Group Behavior*, pages 185–208, 1987.

[76] Hoon Hong Ng. Collective Free Music Improvisation as a Sociocommunicative Endeavor : A Literature Review. *National Association for Music Education*, 2018.

[77] Thomas Berger and Peter L. Luckmann. *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. Double and Company, New York, 1967.

[78] M. Luck and M. d'Inverno. A formal framework for agency and autonomy. In *Proceedings of the First International Conference on Multi-Agent Systems*, pages 254—260. AAAI Press/MIT Press, 1995.

[79] Oliver Bown and Jon McCormack. Creative agency: A clearer goal for artificial life in the arts. In George Kampis, István Karsai, and Eörs Szathmáry, editors, *ECAL (2)*, volume 5778 of *Lecture Notes in Computer Science*, pages 254–261. Springer, 2009.

[80] Michael Cook and Simon Colton. Redesigning Computationally Creative Systems For Continuous Creation. In *Proceedings of the Ninth International Conference on Computational Creativity, ICCC 2018*, 2018.

[81] Vlad Petre Glăveanu. From individual agency to co-agency. In *Constraints of Agency*, pages 245–265. Springer, 2015.

[82] Philip R Cohen and Hector J Levesque. Intention is choice with commitment. *Artificial intelligence*, 42(2-3):213–261, 1990.

[83] Francisco Câmara Pereira. *Creativity and artificial intelligence: a conceptual blending approach*, volume 4. Walter de Gruyter, 2007.

[84] A. Mordvintsev et al. Deepdream-a code example for visualizing neural networks. *Google Research*, 2:5, 2015.

[85] Leon A Gatys, Alexander S Ecker, and Matthias Bethge. Image style transfer using convolutional neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2414–2423, 2016.

[86] Robert R McCrae. Creativity, divergent thinking, and openness to experience. *Journal of personality and social psychology*, 52(6):1258, 1987.

[87] Mark A Runco. *Creativity: Theories and themes: Research, development, and practice*. Elsevier, 2014.

[88] Crystal Gibson et al. Enhanced divergent thinking and creativity in musicians: A behavioral and near-infrared spectroscopy study. *Brain and cognition*, 69(1):162–169, 2009.

[89] Arthur Cropley. In praise of convergent thinking. *Creativity research journal*, 18(3):391–404, 2006.