Rapid and Flexible Creativity in Musical Improvisation: Review and a Model

Psyche Loui a

a Wesleyan University

Address correspondence to: Psyche Loui
Department of Psychology, Wesleyan University
207 High St, Middletown, CT 06459
Email: ploui@wesleyan.edu
Abstract

Creativity has been defined as the ability to produce output that is novel, useful, beneficial, and desired by an audience. But what is musical creativity, and relatedly, to what extent does creativity depend on domain-general or domain-specific neural and cognitive processes? To what extent can musical creativity be taught? To answer these questions from a reductionist scientific approach, we must attempt to isolate the creative process as it pertains to music. Recent work in the neuroscience of creativity has turned to musical improvisation as a window into real-time musical creative process in the brain. Here I provide an overview of recent research in the neuroscience of musical improvisation, especially focusing on multimodal neuroimaging studies. This research informs a model of creativity as a combination of generative and reactive processes that coordinate their functions to give rise to perpetually novel and aesthetically rewarding improvised musical output.

Keywords: creativity; improvisation; entropy; default network; perception action; executive control
The current work seeks to define a rigorous but nuanced model of musical improvisation, by conceptualizing it as a complex system that includes computational, algorithmic, and implementational levels of analysis \(^2\). The mounting research literature suggests that musical improvisation, such as that which is commonly taught in modern jazz training, offers a useful window through which to understand real-time creativity \(^3\). Thus, a model of musical improvisation as a complex system will be informative for cognitive scientists, musical educators, and anyone seeking to better understand creativity.

**Conceptualizing the real-time creative musical process**

Following classic work in cognitive science, a complex system can be described at three levels \(^2\). At the highest, computational level, the model addresses the goal of the overall system: in this case, successful musical improvisation. At the middle, algorithmic level, the model describes the cognitive processes and transformations that must occur to accomplish the goal. And at the lowest, implementational level, the model provides a physical realization of neural substrates necessary for implementing the required cognitive processes.

Musical improvisation lends itself well to scientific study at multiple levels because it involves complex but rapid interactions of many components. In contrast to other forms of musical creativity, such as composition, ideas in musical improvisation (e.g. melodic, harmonic, and rhythmic patterns) are generated and evaluated on a relatively fast time-scale within a performance. Between performances, musical ideas are also generated and evaluated over the course of long-term training in the classroom as well as in private instruction \(^4,5\). Training and experience give rise to the psychological constraints that enable the real-time improvisatory experience. These psychological constraints include the referent (cognitive/perceptual/emotional guidelines or structures), the knowledge base (musical materials and repertoire), and domain-
specific memory for previously encountered auditory-motor patterns. Also guiding improvisations are motor (or biomechanical) constraints which are shaped by experience. The goal of successful improvisations, then, entails filtering the referent through the performer’s own knowledge base to generate fluent, cohesive auditory-motor sequences that are intrinsically rewarding.

How does the cognitive system accomplish this goal? At an algorithmic level, models of creativity entail idea generation and evaluation, in a cognitive cycle akin to the Blind Variation and Selective Retention process that is assessed by psychometric studies such as Divergent Thinking tests. Idea generation is the process of mentally combining or recombining existing elements to give rise to multiple possible solutions, whereas idea evaluation entails selecting from the array of generated ideas, using internally or externally generated feedback. Because feedback can come from multiple sources at different times during or after the performance, this feedforward/feedback cycle between idea generation and evaluation occurs at multiple timescales.

At an implementational level, this interplay of idea generation and evaluation likely entails the coordinated activity of the default mode and executive control networks in the brain as detailed in the next section. As ideas in improvised music are implemented as auditory-motor sequences, the perception and production of these sound targets further engages the auditory perception-action network, which is strengthened in its connectivity by musical training. Figure 1 shows a model of musical improvisation at the computational, algorithmic, and implementation levels.
Figure 1. A model of musical improvisation at computational, algorithmic, and implementation levels. The computational level specifies the goal of real-time musical creativity via improvisation as a system. This is closely tied to the algorithmic level, which describes how the goals specified at the top level are accomplished. At the lowest level are the neural systems that implement the steps of perception and action, idea generation and evaluation, and learning and motor plan selection as shown in the algorithmic level.

A Review of Neuroimaging Studies on Musical Improvisation

Some insights into the neural implementation of musical improvisation come from functional neuroimaging. Several fMRI studies have asked jazz musicians to improvise in the scanner, and compared brain activity or connectivity against control tasks of producing non-improvised sequences (such as musical scales or memorized passages). In the first fMRI study on jazz improvisation, Limb and Braun (2008) compared brain activity in jazz pianists between improvised and overlearned productions of performances of a previously memorized novel
melody. This first study showed increased activity during improvisation in several regions within the frontal lobe including medial prefrontal cortex, cingulate cortex, inferior frontal gyrus and supplementary motor areas, as well as in the auditory processing areas in the temporal lobe including superior and middle temporal gyri. In contrast, the dorsolateral prefrontal cortex was mostly deactivated during improvisation. These differences were not explained by differences in the number or variability of notes played during the improvisation condition, as these were controlled in this study. This pattern of results has given rise to the influential hypothesis that creativity entails an upregulation of mesial prefrontal regions (e.g. medial prefrontal cortex, cingulate cortex) accompanied by a downregulation of lateral prefrontal regions (e.g. dorsolateral prefrontal cortex). However, the specific decrease in DLPFC activity could also arise from the relatively low working memory demands of the improvisation task relative to the control task, which required the recall and production of a newly learned melody. Nevertheless, the balance of mesial to lateral activity can be an important measure, in part because these mesial and lateral prefrontal structures belong to different resting state brain networks, including the default mode network and the executive control network.

Following up on the idea of mesial to lateral activity, Liu et al (2012) investigated functional activity and connectivity using fMRI during freestyle rap, comparing spontaneous lyrical improvisation against conventional, rehearsed performance conditions in freestyle artists. Again, improvisation was associated with increased activity in the medial prefrontal cortex, especially in the left hemisphere, and decreased activity in the dorsolateral prefrontal cortex, especially in the right hemisphere. Furthermore, functional connectivity analyses showed that seed regions in the medial prefrontal cortex were positively associated with the inferior frontal gyrus and the cingulate cortex, and negatively associated with the dorsolateral prefrontal cortex.
and intraparietal sulcus. These results provide further support for the role of dissociated activity between medial and dorsolateral prefrontal cortices in guiding improvisatory behavior. The authors speculate that the medial prefrontal cortex might guide behavior through “alternate cingulate pathways” that effect motor control by “linking intention, affect, language, and action”\(^{18}\). According to this view, the cingulate cortex and the medial prefrontal cortex, although they are anatomically distinct from each other, are nevertheless able to act together\(^ {19}\). The cingulate cortex may serve as a hub that acts upon the auditory perception-action cycle to choose appropriate auditory-motor patterns to maximize reward\(^ {20}\).

The auditory perception-action cycle has been extensively studied due to its importance not only in music, but also in speech and language as well as in hearing more generally. The first cortical waystation of the auditory perception-action pathway lies in the superior temporal lobe, where input from subcortical areas along the auditory pathway is coded in the core, belt, and parabelt areas of the auditory cortex. From the level of the auditory cortex, much evidence supports a dual-stream model of auditory processing. The dorsal stream supports sensorimotor control/integration, whereas the ventral stream supports object-based sound categorization\(^ {21, 22}\). The significance of dorsal vs. ventral pathways in music has also been shown, notably in behavioral and neuroimaging work on tone-deafness, or congenital amusia\(^ {23-25}\). Specifically, the dorsal network involves areas connected by the arcuate fasciculus, which is a major white matter pathway connecting endpoints of cortical grey matter in the superior temporal lobe (superior and middle temporal gyri) and the frontal lobe (inferior frontal gyrus)\(^ {14}\). The ventral network includes middle temporal gyrus and inferior frontal regions connected via the uncinate and inferior longitudinal fasciculi\(^ {14, 26}\). Together, these dorsal and ventral pathways enable sensorimotor translation as well as category-based representation of sound targets in a
feedforward and feedback process. Applied to the study of real-time creativity such as in musical improvisation, this perception-action feedforward-feedback cycle must additionally subserve the generation of novel ideas, a process which must also take into account one’s knowledge base (e.g. previously known melodic fragments or “licks”, or chord progressions).

In another fMRI study, classically trained pianists were asked to improvise on a given melody and produce pseudo-random key-presses, compared to a control task of sight-reading. Both improvisation and pseudo-random conditions showed activity in bilateral inferior frontal gyri and insula, anterior cingulate cortex, left pre-SMA, and bilateral cerebellum. Pseudo-random sequence generation additionally recruited superior frontal gyrus and precentral gyri. The pseudorandom sequence generation task also showed activity in the lingual and fusiform gyri in the occipital lobe. This converges with the Liu et al (2012) and Limb & Braun (2008) findings reviewed above in highlighting the role of mesial and lateral prefrontal cortices, but the differences may have to do with differential task demands, as this is the only study that employed a pseudo-random sequence generation task.

Pinho et al (2014) investigated musical improvisations in jazz and classical pianists and found that while total hours of improvisation experience was negatively associated with activity in the frontoparietal association areas, improvisation training was positively associated with functional connectivity of the bilateral dorsolateral prefrontal cortices, dorsal premotor cortices, and presupplementary motor areas.

Although most studies reviewed thus far showed relatively little activity in inferior frontal cortices, Donnay et al showed that language areas (inferior frontal gyrus) are active during trading fours, which is a form of interpersonal musical interaction common in improvised jazz.
Taken together, fMRI studies of musical improvisation activated frontal, temporal, and parietal areas, with special emphasis paid to a group of prefrontal regions including medial prefrontal and cingulate cortex, dorsolateral prefrontal cortex, and premotor and pre-supplementary and supplementary motor areas. As shown in a recent review \cite{13}, these regions belong to several known functional networks including the default and executive networks.

**Inherent Challenges and Possible Solutions**

Results have generally shown differences in the frontal lobe; specifically the medial prefrontal cortex is frequently active during improvisation whereas the dorsolateral prefrontal cortex is frequently more active during control. However, it is unclear whether the mesial activity reflects over-activation during novel musical idea generation, or whether it reflects under-activation or deactivation during the control condition. Similarly, it is also unclear whether the lateral activity reflects over-activation during control tasks, which often require more memory, or whether it reflects deactivation of the dorsolateral prefrontal cortex during novel musical idea generation.

Besides the above point, there were other discrepant findings between studies, even within the frontal lobe. These discrepancies arise from intrinsic variability in the mental process of improvisation: during a single given moment in the improvisation task, subjects could have been utilizing any number of available mental resources (e.g. visuospatial and/or auditory/phonological components of working memory, autobiographical memory recall, motor planning, attentional selection, and affective communication, just to name a few) to engage in the idea generation and evaluation process. This poses an inherent challenge in task fMRI studies of jazz improvisation.
One approach to address this challenge is to control the stimulus completely by presenting the same predetermined stimuli to all subjects, and to measure the extent to which jazz improvising musicians differ in their perceptual and cognitive processing of matched stimuli. Although this removes the improvisational process from the study, given the appropriate experimental controls, people with different levels of improvisatory training can be reasonably expected to respond differently to the same stimuli as a result of their training.

Another approach around the inherent challenge is to remove the task from the scanner completely, and to compare resting state functional MRI which captures connectivity of the brain without a task at hand. Subjects are simply asked to daydream in the MRI. “Daydreaming” is associated with resting brain activity in the default mode network, which has been tied to idea generation (Christoff et al, 2016). Thus, comparing the default mode and other networks between subjects with different levels of improvisational experience may also offer a window into neural substrates of stimulus-independent thought processes including creativity.

A third way around the inherent challenge is to compare structural differences in the brain, and to relate these differences to measures of musical production outside of the scanner environment, that might elucidate the structural neural mechanisms of idea generation in a more ecologically valid setting. For each of these approaches, a systematic relationship between brain structure or function and improvisatory behavior can only be established after eliminating as many other sources of confounds as possible via careful selection of active control groups.

In the remainder of this article I review a series of recent studies that uses each of these three approaches to tackle the problem of musical improvisation while circumventing the inherent challenge of variability in the improvisational process. In all studies we use multiple control groups, comparing jazz improvising musicians, classical non-improvising musicians, and
non-musicians. Classical and jazz groups are matched on pitch discrimination thresholds, duration of general musical training, and in familiarity with their instrument, but only the jazz group has experience in rapid musical idea generation (for details see Loui et al, 2016; Przysinda et al, 2017). Thus, with the help of multiple control groups we can tease apart whether any differences between groups arise from general perceptual-motor training (by comparing both groups of musicians against non-musicians), or whether they arise from improvisation training per se (by comparing jazz musicians against the other groups).

**Evaluating and Predicting Creativity**

How do we assess jazz musicians’ performance? Here we used an improvisation-continuation task, in which subjects are given a simple, repeated musical motif, and are asked to reproduce and then to improvise on it. The stimulus motifs (https://wesfiles.wesleyan.edu/home/ploui/web/JazzCreativity/ImprovCont/Motives/) and examples of subjects’ recorded productions are available online (https://wesfiles.wesleyan.edu/home/ploui/web/ImprovCont/). Subjective listening to the recordings ensured that all subjects were able to reproduce the stimuli, and also to improvise on them to the best of their ability.

When considering how creative output can be evaluated, it is worth noting that creative works never stand in isolation. Czikszentmihalyi describes creativity as a three-part system that includes the domain (e.g. mathematics; painting), the field (consisting of all experts or professionals in the domain), and the individual creator. The judgment of experts in the field is an important validation of creative output, and the evaluation of musical ideas is crucial to improvisation at the algorithmic level. Thus, we first used a consensual assessment technique to
assess creativity of our subjects’ output. We invited professional musicians and jazz instructors (see Acknowledgements) to listen to each clip and rate them on creativity. Raters showed generally high agreement (r = 0.93; Arkin et al., 2017) and averaged ratings were higher for jazz musicians than for the other two groups.

In addition to subjective methods, we further sought to identify objective, data-driven measures from the subjects’ recorded output that would be useful in predicting experts’ creativity ratings, which could then be applied towards information-theoretic analyses of new recordings. Previous studies of creativity, reviewed above, have used entropy as an information-theoretic measure to analyze their subjects’ behavioral output. Since its first definition, entropy has been used to quantify information content in neuroscience and to model statistical learning in the musical modality (Hansen & Pearce, 2014). Here, we hypothesized that more creative performers would play more notes (i.e. be more fluent) and play more varied notes. We therefore computed two measures, fluency and entropy, for each recording. Fluency was simply defined as the number of notes played per trial. Entropy was defined as the negative sum of the log probability of each note weighted by its probability: \( H(X) = -\sum p_i \log(p_i) \), where \( p_i \) refers to the probability of each note. Intuitively, if one only plays a single note within the whole recording (\( p_i = 1 \)), then \( H(X) \) is 0, whereas if one plays many varied notes, this would result in a positive entropy value. Although this is a simple measure that does not yet take into account any music-theoretical or motoric constraints, this struck us as a valid first-pass measure of creativity, because more creative players could be expected to play more notes and include more different pitches within the course of a single trial. Performances from jazz improvising musicians showed higher fluency and higher entropy. Fluency, entropy, and creativity ratings are all highly correlated (\( r > .8 \)). Fluency and entropy together explain 80% of variance in experts’ creativity.
ratings. Fluency and entropy are highly correlated \((r = .877)\), but fluency explains additional variability in creativity ratings \((\text{partial } r = .49)\) even after accounting for the variability explained by entropy.  

It is worth noting that although entropy is useful as a first measure of the variety of notes played, it cannot be expected to capture all of creativity. Maximum entropy could be achieved by completely random playing on an instrument, whereas maximum fluency would entail playing as many notes as possible, both of which few listeners would find highly creative. Nevertheless, in our sample most subjects were fixated on a few keys, possibly due to the nature of the task, and those who played more notes (high fluency) and more varied notes (high entropy) were also rated as more creative by the experts. Thus, while our current results show positive relationships between fluency, entropy, and creativity, future work is needed to refine the information-theoretic measures that are best applied towards predicting creativity.

Voxel-based morphometry (VBM) was used to relate fluency and entropy to grey matter volume. VBM results showed significant negative associations between entropy and grey matter volume in three regions: the left middle temporal gyrus, the supplementary motor area, and the medial cingulate cortex, whereas fluency was associated with grey matter volume in the left middle temporal gyrus only. These regions correspond to the auditory-motor and default mode networks respectively. Interestingly, all associations observed were negative, with individuals who produced more entropy possessing less grey matter volume in these regions. One hypothetical explanation is that individuals with high grey matter volume in these regions may have had more inhibitory processes leading to less entropic performances.

In addition to grey matter differences, white matter differences were also observed between jazz improvisers and their non-improvising counterparts. A whole-brain Diffusion
Tensor Imaging (DTI) comparison between jazz musicians and controls showed that jazz musicians had higher fractional anisotropy (FA) in mesial regions in the corpus callosum and the cingulum. Furthermore, FA in the middle cingulate cortex was correlated with entropy (but not with fluency). A probabilistic tractography analysis using the mesial significant cluster in the corpus callosum and cingulum as a seed region of interest, and the lateral endpoints of the arcuate fasciculus as waypoint regions of interest, showed higher tract volume and FA in tracts identified between the mesial ROI and the left superior temporal gyrus, and between the mesial ROI and right inferior frontal gyrus. This provides anatomical support for the integration between areas in the lateral perception-action network and mesial areas in the default and executive control networks, which may be related to interhemispheric connectivity in the corpus callosum as well as cognitive control processes in the cingulate cortex.

**Role of Expectation in Idea Evaluation**

Idea evaluation is a crucial aspect of the algorithm in the present model of musical improvisation. The ability to compare and select musical ideas could be assessed by presenting the same musical ideas to multiple groups who differed in their improvisatory experience, and comparing their rapid evaluative responses to the same stimuli. We measured using event-related brain responses to musical chord progressions in jazz improvising musicians, classical musicians, and non-musicians using the well-replicated harmonic expectation paradigm, in which subjects listened to expected, slightly unexpected, and highly unexpected chord progressions, and rated their preference for each chord progression. Behaviorally, jazz musicians preferred the slightly unexpected chord progressions, whereas both other groups preferred the highly expected. Event-related potentials showed larger amplitude of the Early Right Anterior Negativity (ERAN) in response to unexpected chords in jazz musicians, suggesting increased
perceptual sensitivity to unexpected musical events. This ERAN difference was followed by a sharper and higher-amplitude P3b waveform, which indicates more cognitive engagement in jazz musicians \(^{38}\). The P3b was followed by a late parietal positivity (LPP) that was larger in classical musicians compared to jazz musicians, suggesting an acceptance of the unexpected chord on the part of the jazz musicians, but a continued perturbation or delayed return to baseline among the classical musicians. Results highlight the rapid temporal evolution of different types of neural processing of unexpected sounds between classical musicians, jazz musicians, and those with no formal musical training. Notably, the ERAN and P3b correlated with scores on the Divergent Thinking Task \(^{9}\), which is a psychometric test for creativity that does not utilize any musical material. This suggests that the differences in neural processing of unexpected sounds may reflect some domain-general aspects of creativity \(^{38}\).

**Summary and Conclusions**

Creativity is a fundamental capacity of the mind that drives human culture and invention. Despite its importance, creativity has not received the scientific attention it deserves, due to inherent challenges in defining and isolating its component processes \(^{39}\). Precisely because it is hard to define, it behooves us to find a more computationally tractable definition of creativity. Here I outline a model of musical improvisation, a subset of creativity with real-time constraints. I propose that creativity can be redefined as the fluent production of high information content, and that a window into real-time creative behavior is musical improvisation, which can be understood as a complex system with multiple levels. Structural and functional neuroimaging studies highlight the role of mesial and lateral integration in subserving creativity, and the ERP results show that expectation plays a central role. Lateral regions include the endpoints of the
arcuate fasciculus, namely the superior and middle temporal and inferior frontal gyri, which are endpoints of the perception-action pathway. Mesial regions include the cingulate, supplementary motor area, and corpus callosum, which are crucial for interhemispheric communication that facilitate the integration of different functions, as well as medial prefrontal cortex and the cingulate cortex, which have been associated with mind-wandering and cognitive control functions respectively. Future work will try to identify how these pathways are sensitive to training-induced plasticity. Understanding the ability to improvise, and how it can improve as a function of training, may translate to more targeted strategies in music pedagogy, thus having implications for fostering a more creative classroom.

Acknowledgements

Supported by Imagination Institute RFP15-15 and NSF STTR 1720698. I thank Pheeroan akLaff at Wesleyan University, John Baboian from Berklee College of Music, and members of the MIND Lab at Wesleyan University for all their helpful input into this project.

References