



The perception of musical spontaneity in improvised and imitated jazz performances

Annerose Engel* and Peter E. Keller*

Music Cognition and Action Group, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

Edited by:

Josef P. Rauschecker, Georgetown University School of Medicine, USA

Reviewed by:

Kimmo Alho, University of Helsinki, Finland

Amber Leaver, Georgetown University, USA

*Correspondence:

Annerose Engel and Peter E. Keller, Music Cognition and Action Group, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstr. 1a, 04105 Leipzig, Germany.

e-mail: engela@cbs.mpg.de;
keller@cbs.mpg.de

The ability to evaluate spontaneity in human behavior is called upon in the esthetic appreciation of dramatic arts and music. The current study addresses the behavioral and brain mechanisms that mediate the perception of spontaneity in music performance. In a functional magnetic resonance imaging experiment, 22 jazz musicians listened to piano melodies and judged whether they were improvised or imitated. Judgment accuracy (mean 55%; range 44–65%), which was low but above chance, was positively correlated with musical experience and empathy. Analysis of listeners' hemodynamic responses revealed that amygdala activation was stronger for improvisations than imitations. This activation correlated with the variability of performance timing and intensity (loudness) in the melodies, suggesting that the amygdala is involved in the detection of behavioral uncertainty. An analysis based on the subjective classification of melodies according to listeners' judgments revealed that a network including the pre-supplementary motor area, frontal operculum, and anterior insula was most strongly activated for melodies judged to be improvised. This may reflect the increased engagement of an action simulation network when melodic predictions are rendered challenging due to perceived instability in the performer's actions. Taken together, our results suggest that, while certain brain regions in skilled individuals may be generally sensitive to objective cues to spontaneity in human behavior, the ability to evaluate spontaneity accurately depends upon whether an individual's action-related experience and perspective taking skills enable faithful internal simulation of the given behavior.

Keywords: music, improvisation, spontaneity, uncertainty, amygdala, action simulation, human fMRI

INTRODUCTION

Imagine stepping into a jazz club, where you are met by the strains of a pianist negotiating a mesmerizing solo unlike anything that you have heard before. Would you be able to tell from the sounds alone whether the pianist is improvising or playing a rehearsed melody?

Spontaneity is a highly valued quality in many of the world's music performance traditions. Its appreciation by listeners presumably relies upon the interaction of objective auditory cues in a musical performance and the subjective experience and expertise of the listener. It is, therefore, quite likely that individuals differ in their ability to evaluate spontaneity in a performer's actions. This ability, broadly speaking, concerns the sensitivity of one individual to the degree of spontaneity in another's behavior. Such sensitivity is relevant not only to the esthetic appreciation of music and drama, but may be also relevant when inferring others' intentions in everyday situations (e.g., when judging whether someone's behavior is calculated and intended to deceive). Improvised musical performance, however, presents a paradigmatic domain in which to study the perception of spontaneity in human behavior.

Musical improvisation is a creative process during which a performer aims to compose novel music by deciding what sounds to produce – and when, as well as how, to produce them – within the real-time constraints of the performance itself (Pressing, 1988). Fluid melodic invention during improvisation is typically achieved through a mixture of spontaneous decision-making and skills honed through deliberate practice. Improvisers invest effort into developing vocabularies of musical patterns (including pitch and

rhythmic sequences), which they then, during performance, freely combine and vary in a manner that is sensitive to the prevailing musical form and stylistic context (Ashley, 2009). Improvisation thus differs from scripted and rehearsed performance, where the goal is to reproduce or re-interpret previously composed music, albeit often in a manner that is intended to sound spontaneous. Despite noteworthy attempts to elucidate the cognitive underpinnings of improvisation (Pressing, 1988, 1998), the precise nature of the processes that enable a performer to invent melodic material in real-time remain “shrouded in mystery” (Ashley, 2009).

Our own previous research (Keller et al., in press) has sought to identify auditory cues to musical spontaneity through the analysis of performance timing and intensity in improvised and rehearsed jazz piano solos. In this work, jazz pianists first improvised melodies and later imitated excerpts from their own and others' improvisations. The analysis of event timing (i.e., the duration of intervals between successive keystrokes) and intensity (i.e., the force of each keystroke, which determines loudness) indicated that the entropy of both measures was greater for improvised than imitated melodies (though, interestingly, the effect for timing was not reliable when imitating one's own improvisations in familiar jazz styles). Information theoretical approaches to psychology assume that entropy – a measure of the randomness of a probability distribution of values (Shannon, 1948) – reflects uncertainty in human behavior (e.g., decision-making; see Berlyne, 1957; Koechlin and Hyafil, 2007). In view of this, the finding that timing and intensity are more variable in improvisations than in imitations may indicate

irregularities in force control associated with fluctuations in certainty about upcoming actions – i.e., when spontaneously deciding which keys to strike – during improvised musical performance.

The current study extends this work to behavioral and brain processes associated with the perception of musical spontaneity by listeners. Specifically, in a functional magnetic resonance imaging (fMRI) experiment, we investigated the ability of musically trained listeners to differentiate between excerpts from the improvised and rehearsed jazz piano solos that were analyzed by Keller et al. (in press). Our earlier finding that the entropy of timing and intensity is higher in improvised than imitated melodies suggests that these parameters could potentially provide listeners with reliable cues to musical spontaneity. The main research questions addressed here concern (a) whether skilled listeners would be able to judge accurately whether the performances were improvised or rehearsed, and (b) whether listening to improvisations is associated with patterns of brain activation that are distinct from those associated with listening to rehearsed performances. We expected that musically trained listeners would be able to evaluate spontaneity, and that their ability to do so would be grounded in neural mechanisms that allow them to mentally simulate – and thus predict to a certain degree – the performer's actions. Such simulation should allow the skilled listener to experience, at least partially, the brain states associated with improvised and rehearsed musical performance. Precise hypotheses about the behavioral and brain processes that enable the evaluation of musical spontaneity are formulated below.

With respect to listener behavior, we assumed that the perception of musical spontaneity is based on the detection of auditory cues reflecting uncertainty in the performer, and, therefore, that the ability to judge whether a performance is improvised or rehearsed depends on the listener's sensitivity to fluctuations in parameters such as event timing and intensity. Sensitivity to these fluctuations could potentially be influenced by factors that affect general responsiveness to uncertainty in other individuals' behavior, as well as by factors related more specifically to uncertainty in the production of piano melodies. The former, general factors may include socio-cognitive variables like empathy, i.e., the ability to understand others' feelings (of uncertainty, in the present case). Factors related more specifically to the perception of piano melodies include the listener's own experience at playing the piano. Through such experience, an individual has the opportunity to learn about the effects of uncertainty on the variability of timing and intensity in their own playing. Experienced listeners – especially if highly empathic – may therefore be able to recognize these hallmarks of spontaneity in another pianist's performance.

A distinction between domain general and music specific processes can, likewise, be hypothesized with respect to the brain mechanisms underlying the perception of musical spontaneity. At one level, brain areas that are generally sensitive to behavioral variability related to uncertainty may play a role in the detection of such spontaneity. Studies investigating the neural correlates of the perception of behavioral uncertainty point to the involvement of brain regions including anterior cingulate cortex, insula, and amygdala (Singer et al., 2009; Sarinopoulos et al., 2010). The amygdala is particularly interesting in this regard because it has been implicated in a range of processes that may be relevant to evaluating musical spontaneity, including novelty detection (Wright et al., 2003; Blackford et al.,

2010) and the perception of emotionally neutral stimuli that are ambiguous or difficult to predict in terms of their timing (Hsu et al., 2005; Herry et al., 2007).

In addition to brain regions that are generally sensitive to behavioral uncertainty, neural mechanisms that enable a skilled listener to perceive uncertainty in a performer's actions on a specific instrument (in this case, the piano) may facilitate the evaluation of musical spontaneity. Functional links between perceptual and motor processes constitute such a mechanism. Considerable evidence for such perception–action links has accumulated in the auditory domain (Kohler et al., 2002; Gazzola et al., 2006), notably in the context of music, where auditory–motor associations develop with experience playing an instrument (Bangert et al., 2006; Lahav et al., 2007; Mutschler et al., 2007; Zatorre et al., 2007). Relevant studies have shown that listening to music that belongs to an individual's behavioral repertoire leads to the activation of sensory and motor-related brain areas, including the anterior insular cortex, the frontal operculum (Mutschler et al., 2007), the inferior parietal lobe (IPL), and the ventral premotor area (vPM; Bangert et al., 2006; Lahav et al., 2007). The IPL and vPM (in addition to Brodmann's area 44) have also recently been discussed in connection with a posterodorsal auditory pathway that subserves the processing of speech and music-related communicative signals (Rauschecker and Scott, 2009; Rauschecker, 2011).

The co-activation of sensory and motor areas (in the absence of overt movement) is consistent with the proposal that action perception recruits covert sensorimotor processes that internally simulate the observed action (Rizzolatti and Craighero, 2004). Studies of individual differences suggest that activity in action simulation networks increases with increasing action-specific motor expertise (Bangert et al., 2006; Lahav et al., 2007; Mutschler et al., 2007) and, furthermore, is relatively strong in individuals who score highly on self-report measures that assess dimensions of empathy such as “perspective taking” (Gazzola et al., 2006). Researchers in the field of social cognition have argued that action simulation plays a role in generating online predictions about upcoming events in order to facilitate action perception, action understanding, and the coordination of one's own actions with those of others (Wilson and Knoblich, 2005; Sebanz and Knoblich, 2009; see also Gallese et al., 2004; Schubotz, 2007).

In the context of music listening, internal simulation processes may trigger anticipatory auditory images of upcoming sounds (Keller, 2008; Leaver et al., 2009). Action simulation may thus facilitate music perception by utilizing the skilled listener's motor system to enable the real-time prediction of acoustic parameters including pitch, rhythmic timing, and sound intensity (Schubotz, 2007; Rauschecker, 2011). The proportion of neural resources recruited by brain networks engaged in action simulation generally varies as a function of the degree to which prediction is challenging (Schubotz, 2007; Stadler et al., 2011). Fluctuations in performance timing and intensity related to a performer's uncertainty may therefore be associated with relatively strong activation of such simulation networks in experienced, empathic listeners.

Viewing action simulation in light of the broader claim that perception and action recruit common neural networks (Hommel et al., 2001; Rizzolatti and Craighero, 2004) suggests that listening to improvisations should be associated with patterns of brain activa-

tion that are partially distinct from those associated with listening to rehearsed performances, as studies examining the execution of these two varieties of action have highlighted their differential processing (Bengtsson et al., 2007; Berkowitz and Ansari, 2008; Limb and Braun, 2008). One such study found stronger activation in a network comprising the dorsolateral prefrontal cortex (DLPFC), the pre-supplementary motor area (pre-SMA), and the dorsal premotor area (dPM) when pianists improvised variations on a visually displayed melody compared to when they reproduced their improvisations from memory (Bengtsson et al., 2007). The authors interpret this result in terms of the involvement of these areas in tasks where participants can choose a response “freely.”

While research studies on musical spontaneity are small in number, a relatively large body of related work has been conducted in the broader field of voluntary action control. In this field, actions are classified along a continuum reflecting the degree to which they are controlled internally (i.e., endogenously) by the agent or externally (i.e., exogenously) by environmental cues (Waszak et al., 2005; Haggard, 2008). Musical improvisation is internally controlled to the extent that actions are chosen freely and spontaneously; imitating a melody is relatively externally controlled, as a pre-existing sequence of sounds provides an exogenous cue that constrains the actions of the imitating performer. Empirical investigations of internally and externally controlled actions outside the music domain – e.g., in arbitrary tasks involving freely selected vs. externally cued button presses – have revealed distinct behavioral (timing) and brain (electrophysiological) signatures for the two modes of action (Waszak et al., 2005; Keller et al., 2006). Further studies using fMRI have found that activity in the rostral cingulate zone (RCZ) of the anterior cingulate cortex is associated with action selection (deciding “what” to do) while activity in a sub-region of the superior medial frontal gyrus (SFG) is associated with action timing (deciding “when” to do it; Mueller et al., 2007; Kriehoff et al., 2009). To the extent that these regions subserve free response selection, they may also be differentially involved during the perception of internally controlled musical improvisations vs. externally guided imitations.

Based on the foregoing, we hypothesized that listeners with jazz piano experience would be sensitive to differences in the degree of musical spontaneity in improvised and imitated jazz piano solos. Specifically, the ability to discriminate between these modes of performance should vary as a function of the listener’s amount of musical experience and empathy, to the extent that these factors affect the ability to simulate the performers’ actions (for both improvisation and imitation) and to recognize auditory cues to uncertainty in the timing and intensity of the performances. With respect to the neural correlates of evaluating musical spontaneity, we hypothesized that brain regions that have been implicated in the detection of behavioral uncertainty (anterior cingulate cortex, insula, and amygdala) would be sensitive to uncertainty related fluctuations in performance timing and intensity. Furthermore, we expected that the probability of judging a performance to be improvised would increase with increasing demands placed on brain regions involved in action simulation (vPM, IPL, anterior insula, frontal operculum) due to perceived unpredictability in the performer’s actions. Finally, to the extent that simulations are high in fidelity, differences in brain activation when listening to

improvisations vs. imitations should be observed in similar regions to those found for the production of improvised and rehearsed music (pre-SMA, dPM, DLPFC) and of internally and externally controlled actions (RCZ, SFG).

To take into account the possibility that some of the above-mentioned brain regions may differentiate between improvisations and imitations even when listeners fail to make accurate explicit judgments, we analyzed the fMRI data in accordance with a 2×2 factorial design that incorporated both the objective classification of stimulus melodies (real improvisations/real imitations) and subjective classifications based on listeners’ responses (judged improvised/judged imitated). The main effect contrast for the objective classification was expected to reveal differences in neural processing related to physical differences between improvised and imitated melodies. The main effect contrast for subjective classifications tested the neural bases of listeners’ beliefs, and was expected to be informative about experience-related processes such as action simulation.

MATERIALS AND METHODS

PARTICIPANTS

The analyzed sample of listeners consisted of 22 healthy male jazz musicians (mean age = 24 years; range 19–32 years) who had on average 12.8 years (SD 6.8) of piano playing experience, 6.8 years (SD 4.8) of which involved playing jazz. Piano was the primary instrument for 7 of the participants and the second instrument for 15 participants. The average amount of time spent practicing the piano per day was 1 h (SD 1), with 0.5 h (SD 0.7) focused on jazz. In addition, participants spent on average 2.4 h/week (SD 3.7) playing piano in ensembles, with 1.5 h (SD 3.1) devoted to jazz.

Further details concerning participants’ musical experience are as follows. Thirteen out of the 22 individuals were, or had completed, studying music at the university level (specializing in jazz performance, music education, church music, or music theory). The whole sample ($n = 22$) played two to five instruments (mean = $3.2 \pm$ SD 1.0). Primary instruments were: piano ($n = 7$), saxophone ($n = 4$), guitar ($n = 3$), voice ($n = 2$) and electric bass, double bass, drums, bassoon, trombone, and organ ($n = 1$ for each). Participants’ current total amount of musical activity included an average of 2.7 h/day (SD 2.3) playing their instruments, with 1.5 h (SD 1.9) devoted to jazz. In addition, participants spent on average 7.3 h/week (SD 5.8) playing in musical ensembles, with 4.1 h (SD 5.5) of ensemble play focusing on jazz.

Twenty participants were right-handed and two were left-handed according to the Edinburgh handedness inventory (Oldfield, 1971). The experiment was performed in accordance with ethical standards compliant with the declaration of Helsinki and approved by the ethics committee of the University of Leipzig.

STIMULI

The experimental stimulus set included 84 10-s excerpts from piano melodies that had been recorded over novel “backing tracks” representing three contrasting styles found in jazz (swing, bossa nova, blues ballad; see Keller et al., in press). Half of the melodies were spontaneous improvisations played by five highly experienced pianists in the three styles. The other half were rehearsed performances produced by each pianist imitating the improvisation of another pianist. Please

see **Audio S1–S6** in Supplementary Material for sound examples. In addition, 21 10-s excerpts from the backing tracks (7 per style, without improvised or imitated melodies) were used in a baseline condition (which was not utilized in the analyses reported in this article).

STIMULUS GENERATION

The production of stimulus materials proceeded via the following steps. First, chord progressions that are characteristic of the three styles (swing, bossa nova, blues ballad) were composed by a professional jazz pianist/composer (Andrea Keller)¹ for use as the backing tracks. These backing tracks, which included stylistically appropriate harmonic material and bass lines, were performed by the pianist/composer on an electronic keyboard (Clavia Nord Electro 2 73) and recorded in musical instrument digital interface (MIDI) format with Anvil Studio (Willow Software)². The blues backing track consisted of four cycles of a 12-bar chord progression and had a total duration of about 160 s; the bossa track consisted of five cycles of a 16-bar chord progression and had a total duration of around 112 s; the swing track consisted of four cycles of a 16-bar chord progression and had a total duration of 86 s.

Six different pianists (with an average of $11.8 \pm SD 5.8$ years piano experience and, of that, $6.0 \pm SD 4.7$ years jazz piano experience) were recruited to create the stimuli for the present study. However, one pianist (who only had one year experience at playing jazz piano) was judged by authors Annerose Engel and Peter E. Keller to be generally poor at imitating other pianists' improvisations, and all of his performances were discarded. The five pianists who were retained had an average of 11.2 years (SD 6.3) of formal piano training, of which 7.0 years (SD 4.5) involved the development of jazz piano skills, including improvisation. Three of these pianists were professional musicians (a jazz pianist, church musician, and music teacher) and two were highly competent amateur musicians. The pianists practiced jazz piano daily ($1.9 \pm SD 1.9$ h; range 0.5–5 h) and played in jazz bands on a weekly basis (for $3.2 \pm SD 2.2$ h on average; range 1–6 h).

The pianists were asked to improvise melodies over the backing tracks on a digital piano (Yamaha Clavinova, CLP150). Anvil Studio was used to present the backing tracks by playing back the relevant MIDI files on a second, identical digital piano. The backing tracks were unfamiliar to the pianists prior to the improvisation session. Charts showing chord symbols that indicated the harmonic progressions in each backing track were visible during improvisation. Each pianist performed three improvisations per backing track, which were recorded in MIDI format using Anvil Studio. From these improvisations, 30–60 s excerpts (selected for their musical integrity by an experienced music producer and author Peter E. Keller) per pianist/style were transcribed by a professional musician using software for musical transcription and notation (Finale 2005, Coda Music Technology/MakeMusic, Inc.)³.

After 4–12 weeks had elapsed, each of the six pianists returned to the laboratory to imitate the selected excerpts from his or her improvisations (self-imitation) and those produced by two of the other pianists (other imitation). Scanned versions of the transcribed

excerpts were sent to the pianists via email approximately one week prior to these imitation sessions, so that the pianists could learn the notes. During the imitation recording sessions, pianists were instructed to reproduce all audible details of the improvised performances, including the notes and stylistic performance parameters related to timing and expression. Pianists were permitted to listen to the original improvisations – and to practice imitating them – as many times as was needed in order to feel confident in reproducing them, and the transcriptions of the improvised excerpts remained in view during the subsequent recording of imitated performances. Two versions of each excerpt were recorded in MIDI format using Anvil Studio. The original improvisation and backing track could be heard while recording the first (“duet”) version, while the second (“solo”) version was accompanied by only the backing track. Pianists were allowed to record several takes of each imitation, until they indicated to the experimenter that they were satisfied that they had produced the best possible imitation.

In addition to the pianist who was judged to be poor at imitating other pianists' improvisations (see above), another individual was technically unable to imitate one of the bossa improvisations properly. Consequently, 14 improvisations (5 blues, 4 bossa, and 5 swing) and 14 corresponding (solo, other) imitations played by five of the pianists were used as the basis for generating stimuli for the fMRI experiment. These performances and accompanying backing tracks were played back on a digital piano under the control of Anvil Studio, and sound output was recorded as .wav audio files by Logic Pro 8.0.2 (Apple, Inc.)⁴. Three 10-s excerpts were then extracted from the audio files of each improvisation, and three matching excerpts were extracted from each imitation. This resulted in 42 10-s improvisations and 42 matched 10-s imitations with accompanying backing tracks (**Audio S1–S6** in Supplementary Material). In addition, 21 10-s excerpts from the backing tracks (7 per style) were recorded (without improvised or imitated melodies) for use in a baseline condition. All audio files were normalized using Adobe Audition (Adobe Systems, Inc.)⁵ in such a way that they had the same average intensity level while relative changes in intensity within each file were retained. A 500-ms fade in and fade out was applied to each file.

ANALYSIS OF STIMULUS (PIANO PERFORMANCE) PARAMETERS

Several performance parameters were extracted from portions of the MIDI files corresponding to each 10-s stimulus item: (a) identity and number of notes played (i.e., keystrokes produced), (b) the duration of inter-onset intervals (IOIs) between successive keystrokes (IOIs, a measure of performance timing), and (c) the relative intensity of keystrokes within a stimulus item (note that intensity, or loudness, is proportional to the force with which a piano key is struck, and can be measured in “MIDI velocity” in arbitrary units ranging from 1 (soft) to 127 (loud)). These performance parameters were compared across improvisations and imitations using two-tailed *t*-tests for independent samples, since improvised and imitated versions of each melody were played by different pianists.

The number of notes played (mean \pm SD) did not differ significantly between improvised items (27.0 ± 7.6) and imitated items (25.9 ± 7.2), $t(82) = 0.67$; $P = 0.51$ (two-tailed *t*-tests for

¹<http://www.andreakellerpiano.com/>

²<http://www.anvilstudio.com>

³<http://www.finalemusic.com/>

⁴<http://www.apple.com/logicstudio/logicpro/>

⁵<http://www.adobe.com/products/audition/>

independent samples). Inspection of the identity of the notes played in each item revealed that the occasional discrepancies between improvisations and imitations were restricted to the performance of ornaments (e.g., “grace notes”) rather than main melodic notes. This confirms that pianists were highly accurate at melodic imitation (which is not surprising, since they played from transcriptions; see Stimulus Generation).

Performance timing and intensity were analyzed separately in each performed melody item by computing the mean, variance, and Shannon’s information entropy (Shannon, 1948) – a measure of the randomness of a probability distribution of values – of IOIs and MIDI velocity values, respectively (see **Table 1**). For calculating Shannon’s information entropy, probability distributions were constructed first for each item using 400 equally sized 10 ms bins for IOIs and 127 bins for MIDI velocity values. Next, Shannon entropy was calculated for each probability distribution $p(x_i)$ using the formula:

$$H(X) = -\sum_{i=1}^N p(x_i) \ln(p(x_i))$$

Valid entropy values range from 0 to 5.99 for IOIs and from 0 to 4.85 for MIDI velocity. Zero entropy represents perfect order (i.e., no randomness) and large entropy values indicate high randomness in a probability distribution.

Two-tailed *t*-tests for independent samples were run to compare improvisations and imitations on each of the above timing and intensity performance measures. With regard to timing, the mean duration of IOIs between successive keystrokes and the variance of IOIs did not differ significantly between improvised and imitated melodies (mean: $t_{(82)} = -0.61, P = 0.541$; variance: $t_{(82)} = -0.20, P = 0.841$), but the entropy of IOIs was significantly higher in improvisations than in imitations ($t_{(82)} = 2.23, P = 0.028$). Furthermore, while the analysis of intensity (i.e., loudness, as indexed by keystroke velocity) revealed no significant differences in mean intensity between improvisations and imitations ($t_{(82)} = -0.13; P = 0.899$), improvisations were significantly higher than imitations in terms of both variance ($t_{(82)} = 6.66; P = 0.000$)

Table 1 | Average values for the mean, variance, and entropy of keystroke timing and intensity (loudness) in the improvised and imitated melodies.

	Improvisations	Imitations
TIMING		
Mean	376.9 ± 117.5	392.7 ± 118.5
Variance	135051 ± 131281	141390 ± 156586
Entropy	2.90 ± 0.23	2.79 ± 0.23
INTENSITY		
Mean	74.1 ± 8.8	74.3 ± 6.7
Variance	276.2 ± 91.4	149.0 ± 83.6
Entropy	2.96 ± 0.24	2.79 ± 0.25

Timing is indexed by inter-onset intervals between successive keystrokes and intensity (loudness) is indexed by musical instrument digital interface (MIDI) velocity. Units of measurement are: ms and ms² for timing mean and variance, respectively, MIDI velocity and velocity-squared for intensity mean and variance, respectively, Shannon’s H for entropy of timing and intensity. Each value represents the average ± SD across stimulus items.

and entropy ($t_{(82)} = 3.23, P = 0.002$) for this performance parameter. These results confirm that performance timing and intensity were relatively unstable in improvisations.

A brief note on the relationship between variance and entropy is required. Variance and entropy are independent measures of the concentration of values in a probability distribution. These measures are equivalent for normally distributed values: Both variance and entropy are low if there is a single concentration of values around the mean of the distribution. Such a situation may arise with intensity (MIDI velocity) data, for example, if a performer plays at a single intensity level with minor fluctuations. For non-normal distributions, variance is high if entropy is high, but not necessarily vice versa. If there are several concentrations of values that are not contiguous but scattered at intervals throughout a distribution, for instance, then variance will be relatively high while entropy may still be low. Thus, a bimodal distribution with values concentrated in its tails – which could arise if a performer used low and high (but no intermediate) intensities while playing – may have low entropy but high variance. The current study does not aim to address the shapes of the timing and intensity distributions generated while playing (though this may be an interesting topic for future research), and variance and entropy are treated as equally informative, alternative indices of variability in these performance parameters.

EXPERIMENTAL PROCEDURE

The fMRI experiment comprised 84 experimental trials, 21 baseline trials (which were not utilized in the analyses reported in this article), and 21 null events. In experimental trials, participants heard a melody together with the accompanying backing track. The task was to judge whether the melody was improvised or imitated. In baseline trials, participants heard only a backing track excerpt and were required to judge whether or not it was played well. During null events, which occurred after every five trials, participants had no specified task and were instructed to relax.

Improvised and imitated items were presented across experimental trials in randomized order, with the constraint that at least 30 experimental or baseline trials intervened between the presentation of improvised and imitated versions of the same melody. Whether the improvised or imitated version of a melody appeared first in the trial order was varied randomly. No more than three improvisations or imitations could appear in immediate succession. The order of appearance of baseline (backing track) stimuli was randomized and there were at least three experimental trials between baseline trials.

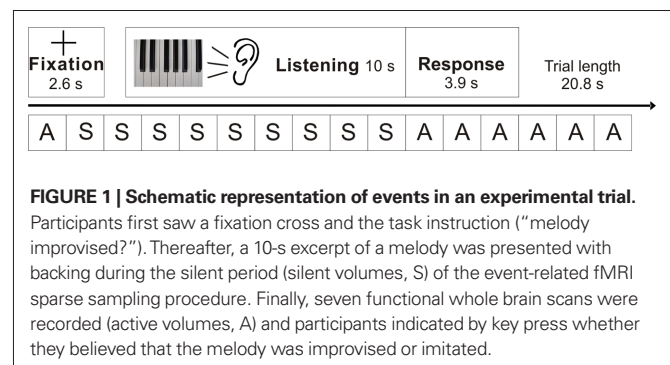


FIGURE 1 | Schematic representation of events in an experimental trial. Participants first saw a fixation cross and the task instruction (“melody improvised?”). Thereafter, a 10-s excerpt of a melody was presented with backing during the silent period (silent volumes, S) of the event-related fMRI sparse sampling procedure. Finally, seven functional whole brain scans were recorded (active volumes, A) and participants indicated by key press whether they believed that the melody was improvised or imitated.

Each trial lasted 20.8 s and consisted of a period without scanner noise (11.7 s) and a period with scanner noise (9.1 s), during which data acquisition took place in a sparse sampling design (see MRI Acquisition and Data Analysis). An example of an experimental trial is depicted in **Figure 1**. Each experimental trial started with a fixation cross (on for 2.6 s) and text informing the participant of the current task: the German equivalent for “melody improvised?”. After a further 400 ms, a 10-s stimulus item was presented, followed by the (re)appearance of the “melody improvised?” text prompt along with text indicating the assignment of response buttons (left or right) to judgment (improvised or imitated). Participants were required to respond by pressing a key with their right index or middle finger before the next trial started (7.8 s later).

Baseline trials had the same procedure as experimental trials, except items consisted only of backing tracks, instructional text at the beginning of a trial was the German equivalent of “backing well played?”, and the participant was required to indicate whether these items were well played (yes or no) by pressing a key (left or right). For null events, the instruction “break” was displayed together with the fixation cross and a black screen was presented after the fixation cross.

During the experiment, participants lay supine on the scanner bed, with the right hand resting on the response box. Visually presented instructions were projected by an LCD projector onto a screen placed behind the participant’s head. The screen was viewed via a mirror on the top of the head coil. All auditory stimuli were presented over scanner compatible headphones (Resonance Technology, Inc.)⁶. Stimulus delivery was controlled by Presentation 11.3 software (Neurobehavioral Systems, Inc.)⁷ running on a computer.

Participants were familiarized with the task prior to the scanning session. They were informed about how the melodies were created (i.e., improvised or imitated) and they were played examples of the backing tracks with and without improvised melodies. In a training phase, participants were presented six 10-s stimulus items, and were asked to judge whether each item was improvised or imitated. Feedback about the correctness of each response was provided and the matched improvised/imitated item was played. Finally, the procedure was practiced (without feedback) first without and then with scanner noise to familiarize participants with the sequence of events: listening to music followed by the presence of scanner noise after the melody presentation (see **Figure 1**). Stimuli presented during practice sessions were not used in the experimental sessions.

After the experiment, participants were debriefed and asked about their strategies for solving the improvised/imitated judgment task. They filled out a questionnaire concerning musical background and the Interpersonal Reactivity Index (Davis, 1980) was administered to assess aspects of empathy.

MRI ACQUISITION AND DATA ANALYSIS

All functional images were collected with a 3T scanner (Medspec 30/100, Bruker, Ettlingen) equipped with a standard birdcage head coil for excitation and signal collection. To avoid contamination by scanner noise during stimulus presentation, we applied a sparse sampling technique, namely interleaved silent steady state echo

planar imaging (Schwarzbauer et al., 2006). During the silent periods of this sequence, the longitudinal magnetization was kept in a steady state by applying silent slice-selective excitation pulses with a repetition time of 1.3 s while data readout was omitted. This procedure avoids T1-related signal decay and therefore ensures that the signal contrast is constant for successive scans. After silent periods, seven volumes were acquired using a 1.3-s repetition time, 30 ms echo time, 68.4° flip angle, and 100 kHz acquisition bandwidth. The matrix acquired was 64 × 64 with a field of view of 19.2 cm, resulting in an in-plane resolution of 3 mm × 3 mm. The slice thickness was 4 mm with an interslice gap of 1 mm. Within each volume, 20 slices were positioned parallel to the anterior and posterior commissure. In total, 882 volumes were collected during one RUN.

Prior to functional image acquisition, two sets of dimensional anatomical images were acquired: T1-weighted modified Driven Equilibrium Fourier Transform images (data matrix 256 × 256, TR = 1300 ms, TI = 650 ms, TE = 10 ms) were obtained with a non-slice-selective inversion pulse followed by a single excitation of each slice; T2* weighted images with the same parameter as the functional scans. After functional image acquisition, geometric distortions were characterized by a B0 field-map scan. The field-map scan consisted of a gradient-echo readout (24 echoes, inter-echo time 0.95 ms) with a standard 2D phase encoding. The B0 field was obtained by a linear fit to the unwrapped phases of all odd echoes.

Structural images were acquired on a 3T scanner (Siemens TRIO, Erlangen) on a different day before the functional scanning session using a T1-weighted 3D MP-RAGE (magnetization-prepared rapid gradient echo) sequence with selective water excitation and linear phase encoding. Magnetization preparation consisted of a non-selective inversion pulse. The following imaging parameters were applied: TI = 650 ms; repetition time of the total sequence cycle, TR = 1300 ms; repetition time of the gradient-echo kernel (snapshot FLASH), TR,A = 10 ms; TE = 3.93 ms; alpha = 10°; bandwidth = 130 Hz/pixel (i.e., 67 kHz total); image matrix = 256 × 240; FOV = 256 mm × 240 mm; slab thickness = 192 mm; 128 partitions; 95% slice resolution; sagittal orientation; spatial resolution = 1 mm × 1 mm × 1.5 mm; 2 acquisitions. To avoid aliasing, oversampling was performed in the read direction (head-foot).

Functional magnetic resonance imaging data were analyzed using SPM5⁸ implemented in Matlab 7.7 (The Mathworks, Inc.)⁹. Images of each participant were pre-processed by realignment to the first image and unwarping was applied. Image distortions were corrected using a field map. Functional images were coregistered to the 3D anatomical image of the participant. The 3D anatomical image was normalized to a Montreal Neurological Institute (MNI) brain template and obtained parameters were used for normalization of the functional data. The voxel dimensions of each reconstructed functional scan were 3 mm × 3 mm × 3 mm. Finally, functional images were spatially smoothed with an 8-mm full-width half-maximum Gaussian filter. A high-pass filter (time constant 128 s) to remove low frequency noise and a correction for autocorrelation (AR(1)) were applied.

In the first level analysis, pre-processed images of each participant were analyzed with a General Linear Model comprising five predictors modeled using a finite impulse response function

⁶<http://www.mrvideo.com/>

⁷<http://www.neurobs.com/>

⁸www.fil.ion.ucl.ac.uk/spm

⁹www.mathworks.com

with the onset of the first acquired volume after the silent period and with a length of 3.9 s (corresponding to the length of three volumes). It was assumed that the first three volumes acquired in each trial mainly reflect brain activity associated with listening to stimulus items rather than activity associated with the following motor response (given a 4- to 6-s lag of the hemodynamic response). Stimuli were assigned to the predictors according to whether they were improvised or imitated (objective classification) and whether they were judged to be improvised or imitated (subjective classification). Resulting predictors cover (1) improvised melodies judged to be improvised; (2) improvised melodies judged to be imitated; (3) imitated melodies judged to be improvised; (4) imitated melodies judged to be imitated; and (5) backing tracks without melody. Contrasts for predictors 1–4 of the first level analysis for each individual participant were entered into a one way analysis of variance (ANOVA) model for a second-level group analysis. Results for contrasts addressing the baseline (backing track) condition are not reported in this article.

Additionally, two parametric analyses were conducted. In the first level of these analyses, activity associated with listening to a melody (regardless of whether it was improvised or imitated and irrespective of participants' judgments) was modeled by a single predictor. Each item of that predictor was weighted according to its entropy of timing or entropy of intensity value (see Analysis of Stimulus (Piano Performance) Parameters). Contrast images for the parameter were used in a second-level one-sample *t*-test for random effects analyses. Significant activations in this analysis reflect correlations between brain activity and the given performance parameter (i.e., entropy of timing or intensity).

The following regions of interest (ROIs) were chosen to test our prior hypotheses based on the literature reviewed in the Introduction: amygdala, vPM, IPL, anterior insula, frontal operculum, pre-SMA, dPM, DLPFC, RCZ, and SFG. For these areas, we report activations that were significant at the $P < 0.05$ level, corrected for multiple comparisons (using the familywise error rate, FWE) for anatomically defined regions built with the WFU PickAtlas¹⁰. For the RCZ and the SFG we used a 10-mm sphere centered on the MNI coordinates ([6, 20, 40] and [-18, 10, 59],

respectively), reported by Kriehoff et al. (2009). Additionally, we report the results of a whole brain analysis at a more liberal significance level of $P < 0.001$, uncorrected for multiple comparisons. To illustrate the results, the parameter estimates of activation clusters were extracted and pooled using MarsBaR¹¹.

RESULTS

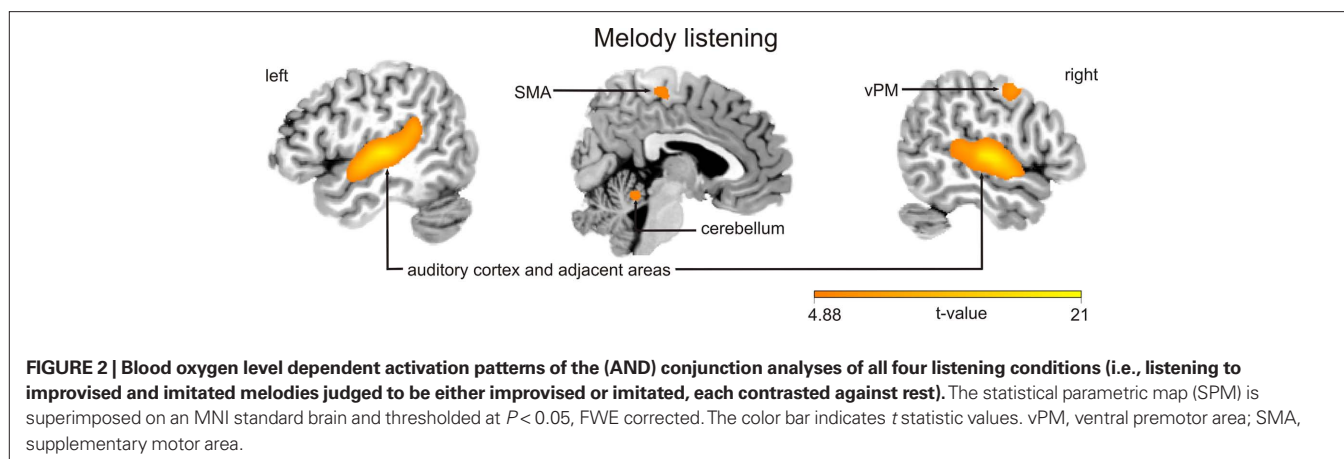
BEHAVIORAL DATA

Listeners were able to judge whether a melody was improvised or imitated with an average correct response rate of 55% (SD 5.4), which is significantly better than chance (50%; two-tailed one-sample *t*-test $t_{(21)} = 4.2$; $P = 0.000$). Although participants were not very confident in their judgments (the mean rating on a scale ranging from 1 = "seldom sure" to 4 = "most of the time sure" was 1.95; SD 0.84), correct response rates covered a fairly wide range: 60–65%, five participants; 55–59%, eight participants; 50–54%, five participants; 44–49%, four participants. As hypothesized, systematic relations were observed between judgment accuracy and self-report measures of musical experience, indicating that accuracy was negatively correlated with the age at which individuals started playing the piano (Pearson's $r = -0.40$, one-tailed $P = 0.033$), particularly jazz piano ($r = -0.48$, one-tailed $P = 0.023$), and positively correlated with the number of hours per week spent playing jazz piano in ensembles ($r = 0.40$, one-tailed $P = 0.037$). The accuracy of improvised/imitated judgments was also positively correlated with scores on the "perspective taking" subscale (addressing the tendency to spontaneously adopt the psychological point of view of others) of a questionnaire (Davis, 1980) assessing dimensions of empathy ($r = 0.44$, one-tailed $P = 0.020$).

An item analysis was conducted to examine the relationship between subjective judgments for each stimulus item (averaged across listeners) and objective measures of performance instability (i.e., variance and entropy of keystroke timing and intensity; see Materials and Methods). This analysis revealed that the probability of judging an item to be improvised was positively correlated with the entropy of timing ($r = 0.31$, two-tailed $P = 0.004$) and the variance of intensity ($r = 0.40$, two-tailed $P = 0.000$). These findings corroborate listeners' self-reports about their task strategies. During

¹⁰<http://fmri.wfubmc.edu/cms/software#PickAtlas>

¹¹<http://marsbar.sourceforge.net>



debriefing, many listeners reported using information about performance timing and rhythm ($n = 16$ out of 22 listeners) and intensity variations ($n = 12$) as a basis for judgments. Additional factors and strategies mentioned by listeners included intuition

($n = 4$), phrasing ($n = 1$), the “interaction” between backing and melody ($n = 3$), and imagining how they themselves would play the melody ($n = 3$).

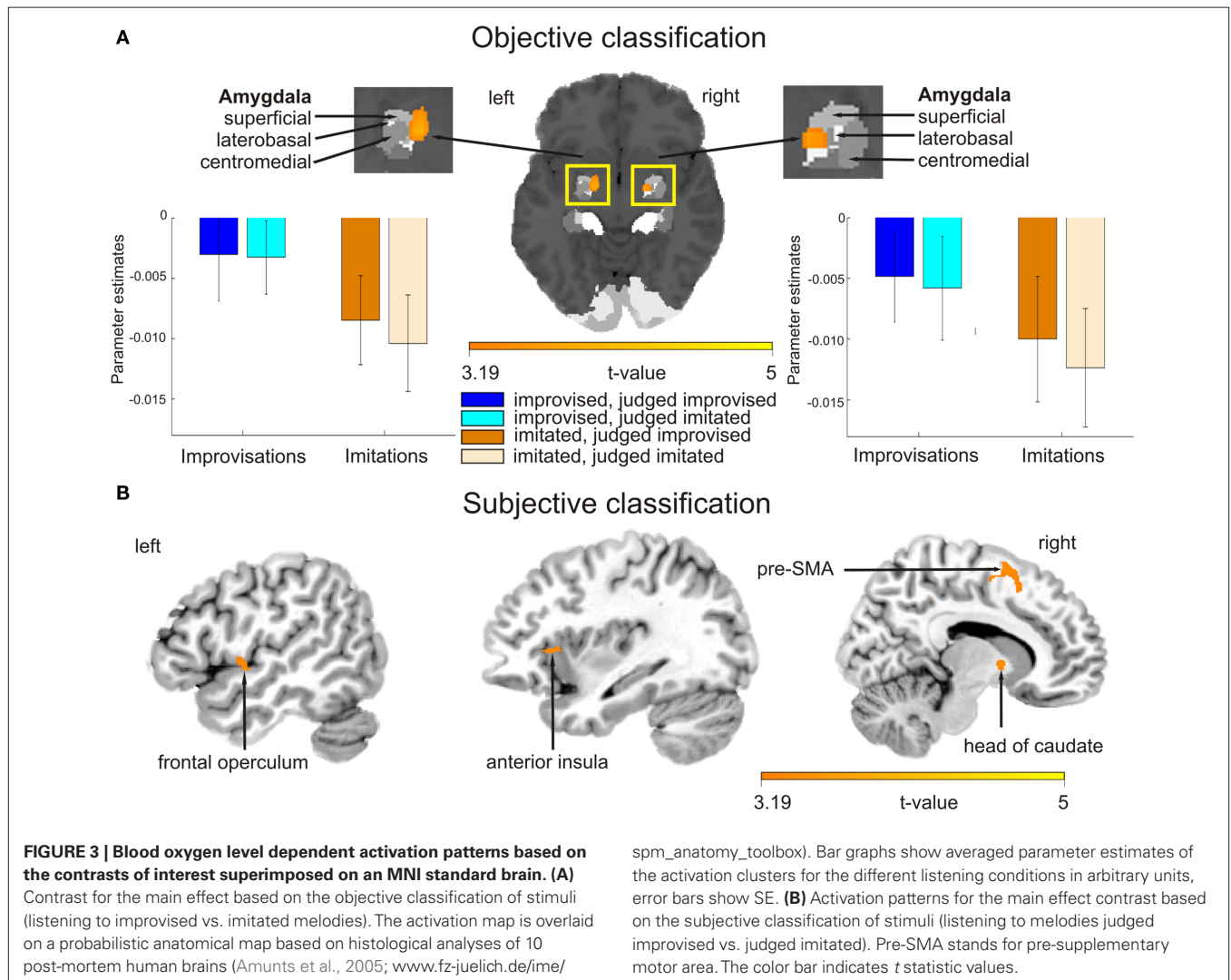
Table 2 | Common brain activation during the perception of improvised and imitated melodies (conjunction analysis).

Anatomical region	Hemisphere	MNI coordinates			Z-score
		x	y	z	
Auditory cortex	R	57	-15	3	>10
	L	-48	-21	3	>10
Precentral gyrus (vPM)	R	57	0	42	7.18
Cerebellum	R	3	-45	-15	5.57
Supplementary motor area	R	6	-24	60	5.56

The values shown are Montreal Neurological Institute (MNI) coordinates for significant activation maxima of clusters in the random effects analyses. vPM, ventral premotor area; R, right; L, left.

FMRI DATA

Functional magnetic resonance imaging data were analyzed according to a 2×2 factorial design that took into account the objective classification of stimuli (real improvisations/real imitations) and subjective classifications based on participants' responses (judged improvised/judged imitated). First, a logical “AND” conjunction analysis (see **Figure 2**; **Table 2**) examining hemodynamic responses common to all four stimulus classifications (improvised and imitated melodies, judged to be improvised or imitated; each contrasted against rest) revealed significant blood oxygen level dependent (BOLD) signal changes bilaterally within the primary and the secondary auditory cortices. These activations extended to several other regions including the superior and middle temporal gyrus, the rolandic operculum, the insula, parietal areas, and various subcortical areas, such as the caudate nucleus, and hippocampal and parahippocampal regions



of the left hemisphere. Further activity occurred in motor-related areas, including the cerebellum, supplementary motor area (SMA; mainly in SMA proper) extending into primary motor area 4a, and the right precentral gyrus (vPM).

The main effect contrast for the objective classification of stimuli (listening to real improvisations vs. real imitations) revealed a bilateral activation cluster in the amygdala nuclei region (left amygdala, $P = 0.010$; right amygdala, $P = 0.019$). To localize this activation more precisely, the coordinates of these activation clusters were compared with anatomical probability maps (Amunts et al., 2005). This comparison revealed overlap mainly with the superficial group, i.e., the cortical part, of the amygdala (see **Figure 3A**; **Table 3**). The parameter estimates displayed in **Figure 3A** show that improvised melodies elicited more activity in the amygdala than imitated melodies regardless of listeners' judgments. No further activations were found in the other ROIs or in a whole brain analysis at a significance level of $P < 0.001$ (uncorrected).

The reverse objective contrast comparing activity associated with listening to real imitations vs. listening to real improvisations yielded no significant differences in any ROIs. In a whole brain analysis, this contrast showed two small activations in the left hippocampal gyrus (MNI coordinates: $x = -45, y = -75, z = 15$; $Z = 3.30$) and the left middle occipital gyrus (MNI coordinates: $x = -33, y = -21, z = -21$; $Z = 3.69$) at a significance level of $P < 0.001$ (uncorrected).

The main effect contrast for the subjective classification of stimuli (listening to melodies judged to be improvised vs. melodies judged to be imitated) revealed differential activity in several ROIs (see **Figure 3B**; **Table 3**): pre-SMA ($P = 0.005$), left anterior insula ($P = 0.046$), left frontal operculum ($P = 0.050$), and RCZ ($P = 0.024$). No differences in activation were found in the left SFG at the coordinate reported by Kriehoff et al. (2009). However, the activation cluster in pre-SMA extended to the SFG in the right

hemisphere. No differences in activation for judged improvisations vs. judged imitations were found in the other ROIs. In addition to the expected activations, a whole brain analysis (significance level $P < 0.001$, uncorrected) revealed further activation clusters on the right head of the caudate nucleus and on the right middle frontal gyrus.

The reverse subjective contrast, which compared brain activity associated with listening to melodies that were judged to be imitated vs. listening to melodies that were judged to be improvised, yielded no significant differences in any ROIs or the whole brain analysis at a significance level of $P < 0.001$ (uncorrected).

Finally, two sets of parametric analyses were conducted to examine relationships between observed brain activations and features of the stimuli pertaining to timing and intensity in the musical performances. In these analyses, each stimulus item (i.e., each melody regardless of its true status as improvised or imitated) was weighted according to either the entropy of IOIs (timing) or the entropy of keystroke velocities (intensity). Results indicated that the entropy of both measures was correlated positively with activity in the left amygdala ($P = 0.003$ for timing; $P = 0.001$ for intensity; see **Table 4A**; see **Figure 4** for entropy of timing results). That is, higher entropy in performance timing and intensity was associated with stronger activity in the left amygdala. Additional positive correlations ($P < 0.001$, uncorrected) were found in the left and right auditory cortex for entropy of timing and intensity, and entropy of timing was correlated positively with activity in the periaqueductal gray of the brainstem. Separate parametric analyses that examined improvised items only yielded similar results at a lower level of statistical significance ($P < 0.005$, uncorrected; see **Table 4B**).

DISCUSSION

The current study addressed the perception of musical spontaneity by examining differences in brain activation associated with listening to improvised vs. imitated jazz piano performances. Behavioral data indicated that listeners (experienced jazz musicians) were able, on average, to classify these performances as improvisations or imitations at an accuracy level that – despite being low (55%) – was significantly better than chance. Listeners' judgments quite likely reflected their sensitivity to differences in the variability of timing and intensity in improvised and imitated performances. Analyses of the performances themselves revealed that the entropy of keystroke timing and intensity was generally higher during improvisation than imitation. This suggests that the spontaneous variability of motor control parameters governing pianists' movement timing and finger force was greater when inventing melodies than when producing rehearsed versions of these melodies. It may be the case that a performer's degree of (un)certainly about upcoming actions fluctuates more widely during improvisation than imitation (see Keller et al., in press). Evidence that listeners used these performance parameters as cues to spontaneity was provided by the finding that performances were likely to attract "improvised" judgments to the extent that they were characterized by variable timing and intensity. Statements made by listeners during post-scan interviews corroborated the notion that judgments were often based on these cues.

Table 3 | Differences in brain activation based on the objective and subjective classification of stimuli.

Anatomical region	Hemisphere	MNI coordinates			Z-score
		x	y	z	
OBJECTIVE CLASSIFICATION: IMPROVISATION > IMITATION					
Amygdala (mainly superficial group)	L	-15	-6	-9	3.80
	R	18	-6	-9	3.60
SUBJECTIVE CLASSIFICATION: JUDGED IMPROVISED > JUDGED IMITATED					
Pre-SMA	R	15	6	60	3.91
RCZ	R	9	15	45	3.33
Insula	L	-36	24	3	3.28
Frontal operculum	L	-54	0	3	3.21
Medial frontal gyrus	R	36	3	54	3.45
Head of caudate	R	9	6	0	3.84

The values shown are Montreal Neurological Institute (MNI) coordinates for activation maxima of clusters. SMA, supplementary motor area; RCZ, rostral cingulate zone; R, right; L, left.

Table 4 | Activations as a function of the entropy of timing and intensity while listening to stimuli. Results are shown for parametric analyses of (A) improvisations and imitations pooled and (B) improvisations only.

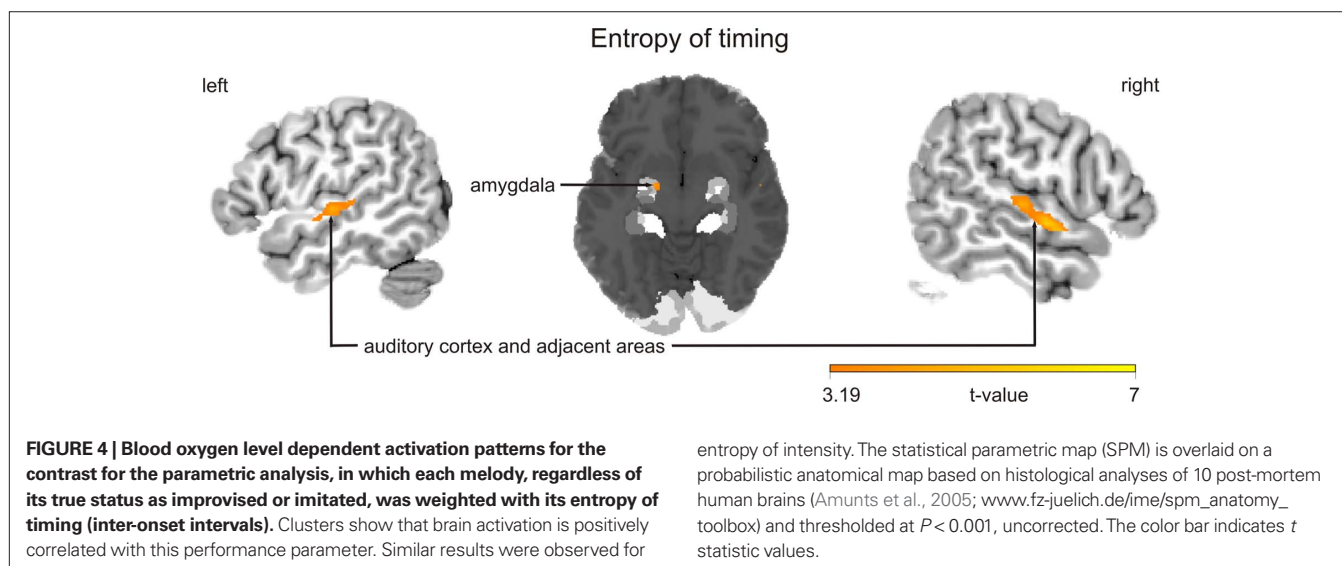
Anatomical region	Hemisphere	MNI coordinates			Z-score
		x	y	z	
A: PARAMETRIC ANALYSIS OF POOLED IMPROVISATIONS AND IMITATIONS					
Entropy of timing					
Amygdala	L	-15	-3	-9	4.19
Auditory cortex	R	45	-21	6	3.65
	L	-42	-30	9	4.73
Brainstem, periaqueductal gray	L/R	3	-30	-3	3.84
Entropy of intensity (loudness)					
Amygdala	L	-15	-3	-9	4.43
Auditory cortex	R	54	-12	3	4.56
	L	-42	-24	6	4.01
B: PARAMETRIC ANALYSIS OF IMPROVISATIONS ONLY					
Entropy of timing					
Amygdala	L	-18	-3	-9	2.69
Auditory cortex	R	48	-21	6	3.23
	L	-42	-27	9	3.77
Brainstem, periaqueductal gray	L/R	6	-33	-6	3.83
Entropy of intensity (loudness)					
Amygdala	L	-15	-3	-9	2.70
Auditory cortex	R	51	-9	0	4.58
	L	-42	-33	9	3.95

The values shown are Montreal Neurological Institute (MNI) coordinates for activation maxima of clusters in (A) and matching activation maxima in (B). R, right; L, left.

The examination of individual differences in the ability to make accurate improvisation/imitation judgments revealed that accuracy was positively correlated with listeners' musical experience and scores on a self-report measure of the "perspective taking" dimension of empathy. These factors may have influenced the detection of variability in timing and intensity; indeed, previous research has shown that musical training enhances auditory sensitivity to timing deviations (Rammsayer and Altenmueller, 2006) and intensity variations in piano tones (Repp, 1995). Musical experience and the ability to "put oneself in another's shoes" may also have enhanced the listener's ability to internally simulate nuances in movement-related activity and feedback that the performer experienced during improvisation and imitation. Broadly consistent with this notion, previous work suggests that covert simulations of musical performances are strong to the extent that the listener has relevant overt behavioral experience (Lahav et al., 2007; Mutschler et al., 2007) and, furthermore, that individuals who obtain high scores on measures of perspective taking are especially good at simulating others' somatosensory experiences (Gazzola et al., 2006).

Analyses of listeners' hemodynamic responses were conducted on the basis of two contrasts applied to the fMRI data. The first contrast – which was based on the objective classification of stimuli as real improvisations vs. real imitations – revealed that listening to improvisations was associated with relatively strong activity in the amygdala. This structure may play a role in the detection of cues to behavioral uncertainty in physical stimulus parameters such as random fluctuations in performance timing and intensity.

Classical views describing amygdala involvement in threat detection, fear conditioning, and the processing of negatively valenced emotional stimuli (LeDoux, 2000; Davis and Whalen, 2001) have recently been supplemented by accounts of amygdala function in the context of non-aversive events (Sander et al., 2003; Ball et al., 2007; Sergerie et al., 2008). These accounts range from those postulating general functions, such as the detection and appraisal of stimuli that are relevant to an individual's basic, as well as social, goals and needs (Sander et al., 2003; Ousdal et al., 2008), to those identifying more specific functions. The latter are related to findings that amy-



gdala responses are modulated by stimulus ambiguity (Hsu et al., 2005), novelty (Wright et al., 2003; Blackford et al., 2010), temporal unpredictability (Herry et al., 2007), and – in music – the violation of listeners' (harmonic) expectancies (Koelsch et al., 2008).

In functional terms, the amygdala may be involved in heightening vigilance and attention in response to ambiguity in external signals (Whalen, 1998). Our finding that hemodynamic responses in the amygdala were correlated with entropy of timing and intensity in the stimuli is consistent with this view. On this account, amygdala activation grew stronger with the increasing presence of cues to behavioral uncertainty in the musical performances. Thus, in highly experienced listeners, the amygdala may be involved in detecting cues to musical spontaneity. The fact that amygdala activation was not dependent upon listeners' judgments implies sensitivity to subliminal cues. Our results may therefore be seen to complement existing behavioral and electrophysiological evidence for the pre-attentive processing of timing and intensity variations in auditory sequences (see Repp, 2005; Tervaniemi et al., 2006).

The second contrast applied to the fMRI data was based on participants' subjective classifications of stimuli as improvisations or imitations. It thus tested the neural bases of listeners' beliefs about the spontaneity of each performance, and, in doing so, is potentially more informative than the objective contrast when it comes to examining experience-related processes such as action simulation. The subjective contrast revealed that hemodynamic responses in the pre-SMA (extending to the SFG), RCZ, frontal operculum, and anterior insula were stronger when listening to melodies that were ultimately judged to be improvised than for melodies that were judged to be imitated. These findings are consistent with our hypothesis that listening to improvisations vs. imitations would generally be associated with cortical activations that overlap with those observed in studies of internally vs. externally guided action execution and in work on covert action simulation. Notably, studies examining differences associated with producing improvised vs. imitated or pre-learned melodies have also reported stronger activation of the pre-SMA, RZC, and frontal operculum during improvisation (Bengtsson et al., 2007; Berkowitz and Ansari, 2008; Limb and Braun, 2008), and the anterior insula has been found to be involved in overt vocal improvisation (Brown et al., 2004; Kleber et al., 2007). Research on voluntary action more generally has shown that the RCZ and SFG play roles in selecting what to do and when to do it (Krieghoff et al., 2009) and that the anterior insula is more strongly activated for internally selected than externally cued actions (Mueller et al., 2007), even if the action is not actually carried out (Kuhn and Brass, 2009).

The relatively strong activation of the pre-SMA, RCZ, frontal operculum, and anterior insula when listening to performances that were judged to be improvised in our study may reflect the greater engagement of an action simulation network related to free response selection. Converging evidence that listeners engaged in action simulation in the current task was provided by a conjunction analysis examining overlap in brain areas activated by improvised and imitated melodies (either judged to be improvised or imitated). This analysis revealed the involvement of motor-related areas, including the cerebellum, SMA, and the vPM, in addition to primary and secondary auditory regions. This pattern of activations is consistent with those observed in previous studies on

covert simulation during music listening (Griffiths et al., 1999; Bangert et al., 2006; Brown and Martinez, 2007; Lahav et al., 2007; Mutschler et al., 2007; Zatorre et al., 2007), as well as in studies of musical imagery (Halpern and Zatorre, 1999; Zatorre and Halpern, 2005; Leaver et al., 2009).

The differential involvement of an action simulation network specializing in freely selected responses when listening to judged improvisations vs. judged imitations may reflect differences in the degree of effort (i.e., amount of processing) required by the cognitive/motor system to generate online predictions about upcoming events in the performances. This prediction process, which may involve auditory imagery (Keller, 2008), is presumably effortful to the extent that the listener's expectancies are violated by perceived fluctuations in performance parameters such as timing and intensity. Previous work has shown that the pre-SMA and insula respond to varying processing demands associated with musical rhythm (Chen et al., 2008; Geiser et al., 2008; Grahn and Rowe, 2009). Furthermore, activation of the frontal operculum and anterior insula increases in response to uncommon harmonies in musical chord progressions (Koelsch, 2005; Tillmann et al., 2006). More generally, it has been proposed that the anterior insula is sensitive to the degree of uncertainty in predictions about behavioral outcomes (Singer et al., 2009). Action simulation during music listening may commandeer a relatively large proportion of neural resources in the above regions when the degree of expectancy violation is high, and prediction challenging, due to perceived instability in the performer's actions (c.f., Schubotz, 2007; Stadler et al., 2011). Under such conditions, listeners may be more likely to classify a performance as an improvisation.

CONCLUSION

Spontaneously improvised piano melodies are characterized by greater variability in timing and intensity than rehearsed imitations of the same melodies, and highly experienced, empathic listeners can detect these differences more accurately than expected by chance. Distinct patterns of brain activation associated with listening to improvised vs. imitated performances occur at two levels. At one level, differences based on the objective classification of performances reflect a distinction in the way the brain processes improvisations and imitations independently of whether the listener classifies them correctly. The amygdala seems to be involved in this differentiation, operating as detector of cues to behavioral uncertainty on the part of the performer who recorded the melody. At the other level, differences in brain activation related to the listener's subjective belief that a performance is improvised or imitated were observed. A cortical network involved in generating online predictions via covert action simulation may mediate judgments about whether a melody is improvised or imitated, perhaps based on the degree of expectancy violation produced by perceived fluctuations in performance stability. It should be noted that the above effects were found with musically trained listeners. Whether they generalize to untrained individuals remains to be seen.

The current findings point to a bipartite answer to the question posed at the opening of this article: Although your amygdala may be sensitive to whether the mesmerizing pianist is engaged in spontaneous improvisation or rehearsed imitation, the ability to judge this would depend on whether your musical experience and

perspective taking skills enable faithful internal simulation of the performance. Thus, while certain brain regions may be generally sensitive to cues to behavioral spontaneity, the conscious evaluation of spontaneity may rely upon action-relevant experience and personality characteristics related to empathy.

ACKNOWLEDGMENTS

This research was supported by the Max Planck Society. We thank Andrea Keller for composing and recording the backing tracks used in this study. We are also grateful to Andreas Weber for recording and editing the pianists' improvisations and imitations for use as stimuli, and for assisting with fMRI data acquisition. Finally, we thank Jöran Lepsien for helpful comments on the design, Toralf Mildner for implementation of the fMRI ISSS sequence, Karsten Müller for discussions concerning the ISSS data analysis, and Johannes Stelzer for assisting with the analysis of the MIDI performance data.

REFERENCES

- Amunts, K., Kedo, O., Kindler, M., Pieperhoff, P., Mohlberg, H., Shah, N. J., Habel, U., Schneider, F., and Zilles, K. (2005). Cytoarchitectonic mapping of the human amygdala, hippocampal region and entorhinal cortex: intersubject variability and probability maps. *Anat. Embryol.* 210, 343–352.
- Ashley, R. (2009). "Musical improvisation," in *Oxford Handbook of Music Psychology*, eds S. Hallam, I. Cross, and M. Thaut (Oxford: Oxford University Press), 413–420.
- Ball, T., Rahm, B., Eickhoff, S. B., Schulze-Bonhage, A., Speck, O., and Mutschler, I. (2007). Response properties of human amygdala subregions: evidence based on functional MRI combined with probabilistic anatomical maps. *PLoS ONE*, e307. doi: 10.1371/journal.pone.0000307
- Bangert, M., Peschel, T., Schlaug, G., Rotte, M., Drescher, D., Hinrichs, H., Heinze, H. J., and Altenmüller, E. (2006). Shared networks for auditory and motor processing in professional pianists: evidence from fMRI conjunction. *Neuroimage* 30, 917–926.
- Bengtsson, S. L., Csikszentmihalyi, M., and Ullen, F. (2007). Cortical regions involved in the generation of musical structures during improvisation in pianists. *J. Cogn. Neurosci.* 19, 830–842.
- Berkowitz, A. L., and Ansari, D. (2008). Generation of novel motor sequences: the neural correlates of musical improvisation. *Neuroimage* 41, 535–543.
- Berlyne, D. E. (1957). Uncertainty and conflict: a point of contact between information-theory and behavior-theory concepts. *Psychol. Rev.* 64, 329–339.
- Blackford, J. U., Buckholtz, J. W., Avery, S. N., and Zald, D. H. (2010). A unique role for the human amygdala in novelty detection. *Neuroimage* 50, 1188–1193.
- Brown, S., and Martinez, M. J. (2007). Activation of premotor vocal areas during musical discrimination. *Brain Cogn.* 63, 59–69.
- Brown, S., Martinez, M. J., Hodges, D. A., Fox, P. T., and Parsons, L. M. (2004). The song system of the human brain. *Brain Res. Cogn. Brain Res.* 20, 363–375.
- Chen, J. L., Penhune, V. B., and Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cereb. Cortex* 18, 2844–2854.
- Davis, M., and Whalen, P. J. (2001). The amygdala: vigilance and emotion. *Mol. Psychiatry* 6, 13–34.
- Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog Sel. Doc. Psychol.* 10, 85.
- Gallese, V., Keysers, C., and Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends Cogn. Sci.* 8, 396–403.
- Gazzola, V., Aziz-Zadeh, L., and Keysers, C. (2006). Empathy and the somatotopic auditory mirror system in humans. *Curr. Biol.* 16, 1824–1829.
- Geiser, E., Zaehle, T., Jancke, L., and Meyer, M. (2008). The neural correlate of speech rhythm as evidenced by metrical speech processing. *J. Cogn. Neurosci.* 20, 541–552.
- Grahn, J. A., and Rowe, J. B. (2009). Feeling the beat: premotor and striatal interactions in musicians and nonmusicians during beat perception. *J. Neurosci.* 29, 7540–7548.
- Griffiths, T. D., Johnsrude, I., Dean, J. L., and Green, G. G. (1999). A common neural substrate for the analysis of pitch and duration pattern in segmented sound? *Neuroreport* 10, 3825–3830.
- Haggard, P. (2008). Human volition: towards a neuroscience of will. *Nat. Rev. Neurosci.* 9, 934–946.
- Halpern, A. R., and Zatorre, R. J. (1999). When that tune runs through your head: a PET investigation of auditory imagery for familiar melodies. *Cereb. Cortex* 9, 697–704.
- Herry, C., Bach, D. R., Esposito, F., Di Salle, F., Perrig, W. J., Scheffler, K., Luthi, A., and Seifritz, E. (2007). Processing of temporal unpredictability in human and animal amygdala. *J. Neurosci.* 27, 5958–5966.
- Hommel, B., Musseler, J., Aschersleben, G., and Prinz, W. (2001). The theory of event coding (TEC): a framework for perception and action planning. *Behav. Brain Sci.* 24, 849–878; discussion 878–937.
- Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., and Camerer, C. F. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science* 310, 1680–1683.
- Keller, P. E. (2008). "Joint action in music performance," in *Enacting Intersubjectivity: A Cognitive and Social Perspective to the Study of Interactions*, eds F. Morganti, A. Carassa, and G. Riva, (Amsterdam: IOS Press), 205–221.
- Keller, P. E., Wascher, E., Prinz, W., Waszak, F., Koch, I., and Rosenbaum, D. A. (2006). Differences between intention-based and stimulus-based actions. *J. Psychophysiol.* 20, 9–20.
- Keller, P. E., Weber, A., and Engel, A. (in press). Practice makes too perfect: fluctuations in loudness indicate spontaneity in musical improvisation. *Music Percept.*
- Kleber, B., Birbaumer, N., Veit, R., Trevorrow, T., and Lotze, M. (2007). Overt and imagined singing of an Italian aria. *Neuroimage* 36, 889–900.
- Koechlin, E., and Hyafil, A. (2007). Anterior prefrontal function and the limits of human decision-making. *Science* 318, 594–598.
- Koelsch, S. (2005). Neural substrates of processing syntax and semantics in music. *Curr. Opin. Neurobiol.* 15, 207–212.
- Koelsch, S., Fritz, T., and Schlaug, G. (2008). Amygdala activity can be modulated by unexpected chord functions during music listening. *Neuroreport* 19, 1815–1819.
- Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., and Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science* 297, 846–848.
- Kriehoff, V., Brass, M., Prinz, W., and Waszak, F. (2009). Dissociating what and when of intentional actions. *Front. Hum. Neurosci.* 3:3. doi: 10.3389/fneuro.09.003.2009
- Kuhn, S., and Brass, M. (2009). When doing nothing is an option: the neural correlates of deciding whether to act or not. *Neuroimage* 46, 1187–1193.
- Lahav, A., Saltzman, E., and Schlaug, G. (2007). Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *J. Neurosci.* 27, 308–314.
- Leaver, A., Van Lare, J. E., Zielinski, B. A., Halpern, A., and Rauschecker, J. P. (2009). Brain activation during anticipation of sound sequences. *J. Neurosci.* 29, 2477–2485.
- LeDoux, J. E. (2000). Emotion circuits in the brain. *Annu. Rev. Neurosci.* 23, 155–184.
- Limb, C. J., and Braun, A. R. (2008). Neural substrates of spontaneous musical performance: an fMRI study of jazz improvisation. *PLoS ONE*, e1679. doi: 10.1371/journal.pone.0001679
- Mueller, V., Brass, M., Waszak, F., and Prinz, W. (2007). The role of the

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/Auditory_Cognitive_Neuroscience/10.3389/fpsyg.2011.00083/abstract

Audio S1 | Example of a 10-s excerpt of an improvised melody played over the swing backing track.

Audio S2 | Example of an imitated version of the improvised swing melody in **Audio S1**.

Audio S3 | Example of a 10-s excerpt of an improvised melody played over the bossa nova backing track.

Audio S4 | Example of an imitated version of the improvised bossa nova melody in **Audio S3**.

Audio S5 | Example of a 10-s excerpt of an improvised melody played over the blues ballad backing track.

Audio S6 | Example of an imitated version of the improvised blues ballad melody in **Audio S5**.

- preSMA and the rostral cingulate zone in internally selected actions. *Neuroimage* 37, 1354–1361.
- Mutschler, I., Schulze-Bonhage, A., Glauche, V., Demandt, E., Speck, O., and Ball, T. (2007). A rapid sound-action association effect in human insular cortex. *PLoS ONE*, e259. doi: 10.1371/journal.pone.0000259
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113.
- Ousdal, O. T., Jensen, J., Server, A., Hariri, A. R., Nakstad, P. H., and Andreassen, O. A. (2008). The human amygdala is involved in general behavioral relevance detection: evidence from an event-related functional magnetic resonance imaging go-nogo task. *Neuroscience* 156, 450–455.
- Pressing, J. (1988). "Improvisation: methods and models," in *Generative Processes in Music*, ed. J. Sloboda (Oxford: Oxford University Press), 129–178.
- Pressing, J. (1998). "Psychological constraints on improvisational expertise and communication," in *In the Course of Performance: Studies in the World of Musical Improvisation*, eds B. Nettl and M. Russell (Chicago: University of Chicago Press), 47–68.
- Rammsayer, T., and Altenmueller, E. (2006). Temporal information processing in musicians and nonmusicians. *Music Percept.* 24, 37–49.
- Rauschecker, J. P. (2011). An expanded role for the dorsal auditory pathway in sensorimotor integration and control. *Hear. Res.* 271, 16–25.
- Rauschecker, J. P., and Scott, S. (2009). Maps and streams in the auditory cortex. *Nat. Neurosci.* 12, 718–724.
- Repp, B. H. (1995). Detectability of duration and intensity increments in melody tones: a partial connection between music perception and performance. *Percept. Psychophys.* 57, 1217–1232.
- Repp, B. H. (2005). Sensorimotor synchronization: a review of the tapping literature. *Psychon. Bull. Rev.* 12, 969–992.
- Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192.
- Sander, D., Grafman, J., and Zalla, T. (2003). The human amygdala: an evolved system for relevance detection. *Rev. Neurosci.* 14, 303–316.
- Sarinopoulos, I., Grupe, D. W., Mackiewicz, K. L., Herrington, J. D., Lor, M., Steege, E. E., and Nitschke, J. B. (2010). Uncertainty during anticipation modulates neural responses to aversion in human insula and amygdala. *Cereb. Cortex* 4, 929–940.
- Schubotz, R. I. (2007). Prediction of external events with our motor system: towards a new framework. *Trends Cogn. Sci.* 11, 211–218.
- Schwarzbauer, C., Davis, M. H., Rodd, J. M., and Johnsrude, I. (2006). Interleaved silent steady state (ISSS) imaging: a new sparse imaging method applied to auditory fMRI. *Neuroimage* 29, 774–782.
- Sebanz, N., and Knoblich, G. (2009). Prediction in joint action: what, when, and where. *Top. Cogn. Sci.* 1, 353–367.
- Sergerie, K., Chochol, C., and Armony, J. L. (2008). The role of the amygdala in emotional processing: a quantitative meta-analysis of functional neuroimaging studies. *Neurosci. Biobehav. Rev.* 32, 811–830.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 379–423, 623–656.
- Singer, T., Critchley, H. D., and Preusschoff, K. (2009). A common role of insula in feelings, empathy and uncertainty. *Trends Cogn. Sci.* 13, 334–340.
- Stadler, W., Schubotz, R. I., von Cramon, D. Y., Springer, A., Graf, M., and Prinz, W. (2011). Predicting and memorizing observed action: differential premotor cortex involvement. *Hum. Brain Mapp.* 32, 677–687.
- Tervaniemi, M., Castaneda, A., Knoll, M., and Uther, M. (2006). Sound processing in amateur musicians and nonmusicians: event-related potential and behavioral indices. *NeuroReport* 17, 1225–1228.
- Tillmann, B., Koelsch, S., Escoffier, N., Bigand, E., Lalitte, P., Friederici, A. D., and von Cramon, D. Y. (2006). Cognitive priming in sung and instrumental music: activation of inferior frontal cortex. *Neuroimage* 31, 1771–1782.
- Waszak, F., Wascher, E., Keller, P., Koch, I., Aschersleben, G., Rosenbaum, D. A., and Prinz, W. (2005). Intention-based and stimulus-based mechanisms in action selection. *Exp. Brain Res.* 162, 346–356.
- Whalen, P. J. (1998). Fear, vigilance and ambiguity: initial neuroimaging studies of the human amygdala. *Curr. Dir. Psychol. Sci.* 7, 177–188.
- Wilson, M., and Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychol. Bull.* 131, 460–473.
- Wright, C. I., Martis, B., Schwartz, C. E., Shin, L. M., Fischer, H. H., McMullin, K., and Rauch, S. L. (2003). Novelty responses and differential effects of order in the amygdala, substantia innominata, and inferior temporal cortex. *Neuroimage* 18, 660–669.
- Zatorre, R. J., Chen, J. L., and Penhune, V. B. (2007). When the brain plays music: auditory-motor interactions in music perception and production. *Nat. Rev. Neurosci.* 8, 547–558.
- Zatorre, R. J., and Halpern, A. R. (2005). Mental concerts: musical imagery and auditory cortex. *Neuron* 47, 9–12.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 03 February 2011; paper pending published: 15 February 2011; accepted: 20 April 2011; published online: 03 May 2011.

Citation: Engel A and Keller PE (2011) The perception of musical spontaneity in improvised and imitated jazz performances. *Front. Psychology* 2:83. doi: 10.3389/fpsyg.2011.00083

This article was submitted to *Frontiers in Auditory Cognitive Neuroscience*, a specialty of *Frontiers in Psychology*.

Copyright © 2011 Engel and Keller. This is an open-access article subject to a non-exclusive license between the authors and Frontiers Media SA, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and other Frontiers conditions are complied with.