Journal: **OPEN MIND**

Coordination and Consonance Between Interacting, Improvising Musicians

anonymous submission

Keywords: joint action, distributed cognition, improvisation, time series modeling, music

Abstract

Joint action (JA) is ubiquitous in our cognitive lives. From basketball teams to teams of surgeons, humans often coordinate with one another to achieve some common goal. Idealized laboratory studies of group behavior have begun to elucidate basic JA mechanisms, but little is understood about how these mechanisms scale up in more sophisticated and open-ended JA that occurs in the wild. We address this gap by examining coordination in a paragon domain for creative joint expression: improvising jazz musicians. Coordination in jazz music subserves an aesthetic goal: the generation of a collective musical expression comprising coherent, highly nuanced musical structure (e.g. rhythm, harmony). In our study, dyads of professional jazz pianists improvised in a "coupled", mutually adaptive condition, and an "overdubbed" condition which precluded mutual adaptation, as occurs in common studio recording practices. Using a model of musical tonality, we quantify the flow of rhythmic and harmonic information between musicians as a function of interaction condition. Our analyses show that mutually adapting dyads achieve greater temporal alignment and produce more consonant harmonies. These musical signatures of coordination were preferred by independent improvisers and naive listeners, who gave higher quality ratings to coupled interactions despite being blind to condition. We present these results and discuss their implications for music technology and JA research more generally.

INTRODUCTION

High-level cognition is often achieved by groups of interacting individuals Knoblich, Butterfill, and Sebanz (2011); Sebanz, Bekkering, and Knoblich (2006). Group behavior in joint action (JA) settings is less dependent on isolated individual efforts and more on the ability to coordinate Goldstone and Gureckis (2009); Hasson, Ghazanfar, Galantucci, Garrod, and Keysers (2012). Insight into the mechanisms underlying successful coordination has important implications for how we understand interpersonal interaction, optimize team performance, and engineer human-like artificial intelligence systems Council et al. (2015); Guimera, Uzzi, Spiro, and Amaral (2005); Rebsamen et al. (2010); D. C. Richardson, Dale, and Kirkham (2007). This study examines coordination in collaboratively improvising jazz musicians. Coordination in jazz music subserves an aesthetic goal: the generation of a collective musical expression, and the expertise of professional jazz musicians lies largely in their ability to coordinate and adapt spontaneously in real-time performance. Professional jazz ensembles thus offer a remarkably sophisticated paragon domain to study the basic properties and limits of our capacity to coordinate with one another.

Humans align their behaviors as they interact Hasson and Frith (2016); Pickering and Garrod (n.d., 2004). We spontaneously entrain periodic motions (e.g. postural sway, walking gait), and such entrainment is predictive of successful interaction and performance on joint tasks Demos, Chaffin, Begosh, Daniels, and Marsh (2012); Paxton and Dale (2013); M. J. Richardson, Marsh, Isenhower, Goodman, and Schmidt (2007); Shockley, Richardson, and Dale (2009); Shockley, Santana, and Fowler (2003). Interlocutors tend to mirror one another's posture, speech prosody and align eye gaze to fixate on the same objects as they interact Garrod and Pickering (2009); Louwerse, Dale, Bard, and Jeuniaux (2012); D. C. Richardson and Dale (2005); D. C. Richardson et al. (2007); D. C. Richardson, Dale, and Tomlinson (n.d.). Alignment occurs at more abstract levels as well. Interlocutors mirror vocabulary and syntactical constructions, and come to share common mental representations for situations under discussion Abney, Paxton, Dale, and Kello (2014); Dale and Spivey (2006); Pickering and Garrod (2004). Past JA research demonstrates that alignment is an important interpersonal mechanism that facilitates joint attention and predictive emulation (of a partner's future actions), and streamlines communication by providing a common representational scheme Garrod and Pickering (2009); Pickering and Garrod (2004); D. C. Richardson et al. (n.d.); Sebanz et al. (2006); Sebanz and Knoblich (2009).

Another issue in JA research is whether group behavior is supported by mutual adaptations (bidirectional coordination) or fixed leader-follower roles (unidirectional coordination). Clearly delineated leader-follower roles appear to support stable coordination in many naturalistic JA domains (e.g. conductor of an orchestra, lead dancer in a salsa pair), and experimental studies have affirmed the utility of unidirectional coordination with respect to particular task constraints and participant expertise levels

Curioni, Vesper, Knoblich, and Sebanz (2019); Noy, Dekel, and Alon (2011); M. J. Richardson et al. (2015). On the other hand, finger tapping studies have shown that dyads achieve greater synchronization when mutually coupled compared to unidirectional conditions Demos, Carter, Wanderley, and Palmer (2017); Konvalinka, Vuust, Roepstorff, and Frith (2010). Rather than adopting leader-follower roles, mutually coupled individuals each adapted their own tapping rates to their partner's previous tapping rates Konvalinka et al. (2010). A similar result has been observed in a simplified experimental adaptation of the "mirror game", which requires dyads to synchronize improvised hand movements with one another. Mutually coupled dyads synchronized more fluidly and generated more dynamic movements compared to dyads that were assigned leader-follower roles Noy et al. (2011).

These findings show that mutual coupling often promotes coordination by supporting robust and flexible behavioral alignment. However, they were obtained in idealized experimental paradigms using greatly simplified behaviors (e.g. synchrony of a tapped pulse), so it is unclear whether and how they generalize to more sophisticated coordinated behavior found in the real world. Naturalistic JA is often open-ended, and requires not just behavioral matching but also *complementary* coordination in service of abstract, functional goals (e.g. operating on a patient, generating ideas in group brainstorming sessions) Hasson and Frith (2016). How does mutual coupling shape coordination in these more complex, naturalistic forms of JA? Does mutual coupling support greater behavioral alignment in under-constrained tasks, where this is no explicit goal of synchronization? Does it support complementary coordination, in service of abstract goals?

In this study we use improvised music as a model domain to explore the effects of mutual coupling in the wild. Conveniently, joint music performance is naturally mediated by organizational structures that constrain ensemble coordination. Orchestras are hierarchically organized with fixed leader-follower roles, whereas free improvising jazz ensembles are typically more characterized by feedback loops of mutual influence Borgo (2005); D'Ausilio et al. (2012). Studio recording practices such as "overdubbing" also constrain coordination by sequentially recording individual musical parts. Ensemble performance research has shown that these underlying patterns of coordination are reflected in the music and movements of ensemble members Hennig (2014); Keller (2014); Rasch (1979), such as small temporal asynchronies of co-performer note onsets Demos et al. (2017); Goebl and Palmer (2009); Keller and Appel (2010), and postural sway couplings Chang, Livingstone, Bosnyak, and Trainor (2017).

Improvised music is of particular interest, because the influence of coordination extends beyond sensorimotor coupling and into the music's formal architecture, which is freely evolving over time in its rhythm, melody, harmony, and texture. We might thus expect underlying coordination patterns to constrain these structural elements, similar to how it constrains sensorimotor coupling in scored music performance. Do mutually coupled improvisers engage in bidirectional coordination at the level of notes and rhythms? If so, does this result in higher quality music? Answering these questions will extend our understanding of JA beyond idealized laboratory tasks and into sophisticated, open-ended coordination that occurs in elite artistic performances. It will also yield direct implications for music technology. Results will reveal repercussions of the popular recording technique of overdubbing, and our quantitative measures of improvised musical coordination can be incorporated into artificial interactive music systems Gillick, Roberts, Engel, Eck, and Bamman (2019); Linson, Dobbyn, Lewis, and Laney (2015) and benefit music pedagogy by automating assessment of ensemble performance.

Despite a paucity of cognitive science research on collective improvisation, some notable efforts have begun. Previous studies have shown that improvised musical coordination is shaped by musical context (e.g. playing with a drone versus a swing backing track), and that experimentally manipulated social attitudes (e.g. dominant, caring) are sonically encoded in improvised musical interactions Aucouturier and Canonne (2017); Walton et al. (2018). These studies lay an important foundation, but they did not experimentally isolate mutual coupling between musicians. Moreover, their analyses did not incorporate music theory, and thus the findings are limited to temporal and acoustic coordination properties, and do not extend to more abstract musical phenomena such as the emergence of tonal structure (i.e. harmony, melody).

In the current study we directly manipulate interaction in co-improvising musicians, and examine how different underlying patterns of coordination constrain the exchange and emergence of rhythmic and tonal information. Professional jazz musicians freely improvised in two duo conditions: a *coupled* condition, in which both pianists improvised simultaneously, and a *one-way* condition, in which a single pianist improvised along with a recording of another pianist (a "ghost partner") from a previous *coupled* duet. Improvisations were completely "free" in the sense that there was no predetermined songform, key signature or tempo; the only instruction was to improvise a compelling piece of music *de novo*, as in an actual performance. These duo conditions provided two naturalistic musical settings to isolate the effects of mutual coupling in freely improvising musicians. Whereas *coupled* duos had the ability to mutually adapt to one another, *one-way* duos were restricted to unidirectional coordination (i.e. because the 'ghost partner' was unresponsive to the live musician), as in the common studio recording technique of overdubbing.

Participants were recorded in isolated MIDI[1] tracks as they improvised in each condition. Time series of two fundamental musical features were extracted and analyzed: onset density and tonal consonance. Onset density indexes overall rhythmic activity level, and has been shown to correlate with listener perception of musical tension Farbood (2012). Tonal consonance refers to how different combinations of notes sound on a continuum from dissonant/unstable to consonant/stable Johnson-Laird, Kang, and Leong (2012), and was operationalized using a previously established model of musical tonality, the Tonal Spiral Array Chew (2005); Chew et al. (2014); Herremans, Chew, et al. (2016). We find that interaction condition systematically altered the coordinated musical behavior of dyads, who were more rhythmically coupled and produced more consonant tonal structure when mutually coupled. These effects were paralleled in the subjective experiences of participants as well as non-musician listeners, who preferred *coupled* duets despite being blind to condition. These results are presented and discussed in terms of their implications for music technology and JA research more generally.

METHODS

Participants

28 professional pianists (25 male, 3 female) from the New York City jazz scene participated in this study. Participant age ranged from 21-37. On average participants had over 22 years experience playing piano (sd=5.2) and 15 years experience improvising (sd=4.6). All participants had extensive experience with free improvisation, and received formal training in piano performance and/or jazz studies at elite conservatories. Participants were recruited by word of mouth, and had no prior experience performing with one another.

[1] Musical Instrument Digital Interface (MIDI) is a format for representing music on a computer. It symbolically represents the pitch, volume and timing (onset and offset) of musical note sequences.

122 individuals participated in the listener study. 101 were undergraduate psychology students from Indiana University without any particular musical background, and 21 (19 male, 2 female) were professional jazz musicians, each with over 10 years of experience as improvising musicians, recruited by word of mouth from the NYC music scene. None of these listeners participated in the initial music-generation stage of the study.

Design and Procedure

Participants played a series of short (4-7 minute) 'free' improvisations, with no accompanying stimuli and no prior musical template or constraints. Other than the suggested time frame, the only instruction was to improvise a compelling piece of music, as in a typical performance setting. Participants were informed of the two interaction conditions, but were not told which condition they were playing in on any given trial (and there was no visual or audible indication of condition, see SI). After each trial, they responded to questionnaires indicating their subjective experience playing in the previous trial in terms of: (1) how easy it was to coordinate with their partner (2) how well coordinated they were with their partner (3) quality of the improvised piece and (4) degree to which they played a leader versus a supporter role.

Each participant played at least 3 duets (trials) in each condition, with the same 'live' partner for every *coupled* duet and the same 'ghost' partner for every *one-way* duet. Conditions were interleaved within participant pairs and counterbalanced across pairs to control for possible order effects. Participants were recorded in isolated MIDI tracks, and individual recordings from *coupled* duets yoked *one-way* duets in subsequent sessions, as depicted in the SI. Altogether 50 *coupled* duets and 86 *one-way* duets were collected; duets had an average duration of 342 seconds (min=108 seconds, max=738 seconds, sd=12 seconds). This dataset will be made publicly available on Github upon acceptance of the paper.

A post-hoc study was conducted with populations of naive listeners and expert jazz musicians. Listeners heard 30-second audio clips randomly sampled from duets in both conditions (audio from each pianist was panned to separate ears). After listening to each clip they were asked to rate (1) their enjoyment of the music (2) how well coordinated they perceived the musicians to be and (3) which musician played more of a leader role. Listeners were also asked to guess which condition a clip came from. Each participant heard complementary yoked sets of *coupled* and *one-way* clips. See SI for full specification of the sequencing design, which controlled for possible order and stereo-panning effects.

Tonal Consonance Measure

Our tonal consonance measure is based on the Tonal Spiral Array model, which has been validated against listener ratings and expert music theory analyses Chew (2005); Chew et al. (2014); Herremans et al. (2016). Table 1 shows model ratings for exemplar pitch sets. See SI for specification of the measure.

Table 1. Consonance ratings of exemplar pitch sets.

Pitch Set	Consonance
{C,E,G} (Cmaj)	.65
{C,Eb,G} (Cmin)	.65
{C,B,G}	.54
{C,E,G,F,A,C} (Cmaj + Fmaj)	.49
{C,B}	.48
{C,E,G,F#,A#,C#} (Cmaj + F#maj)	.13
serial (all 12 pitches)	.09

Data Analysis

Listener ratings were analyzed with Bayesian mixed-effects models for each response type, using the brms package in R Bürkner et al. (2017). Instead of predicting enjoyment and coordination ratings directly, models predicted the difference between ratings of *coupled* audio clips minus ratings of correspondingly yoked *one-way* clips, such that positive intercepts indicated preference for *coupled* clips. Leadership ratings within *one-way* trials were modeled such that positive intercepts indicated perception of "ghosts" leading, and negative values indicated perception of live musicians leading. Accuracy of condition guesses was modeled as binomial outcome: whether or not listeners guessed the correct condition, such that positive intercepts indicated above-chance predictions. Models included a predictor for subject type (naive listener or professional jazz musician), and random intercepts per individual.

Bayesian mixed-effects models were also used to analyze time series measures of musical coordination (cross-correlation of onset density and lagged consonance, see Results). Dependent measures were predicted by a fixed-effect of interaction condition, with random intercepts for yoked groupings at the duo and duet levels [2] Unidirectional coordination in one-way duos was analyzed by predicting dependent measures as a function of lag direction (i.e. ghost-to-live versus live-to-ghost), with random-effects for each duo and duet.

RESULTS

Subjective Ratings

Despite being blind to condition, performers and naive listeners both exhibited a strong preference for *coupled* over *one-way* duets. Performers rated *coupled* trials as producing higher quality music (21 out of 26 performers rated coupled higher; probability of success = 0.81; exact binomial test $p<.01$). Coupled trials were also rated as being better coordinated (23 out of 26 performers rated coupled trials as being better coordinated; probability of success = 0.88; binomial test $p<.01$), and more easily coordinated (24 out of 26 performers found it easier to coordinate with their partner on coupled trials; probability .92; $p<.01$). Performers also rated themselves as playing more of a supportive (versus lead) role in *one-way* duos, whereas leadership was rated to be more evenly distributed throughout coupled duos (difference between average ratings within participant by condition; paired-$t(25)=3.16$, $p<.01$).

Bayesian mixed-effects models predicting the difference in listener ratings between coupled clips and correspondingly yoked one-way clips indicated that listeners found coupled clips to be more enjoyable (M = 0.24, SD = .08, 95% CI = [.08, 0.40]) and better coordinated (M=.43,SD=.11,95%CI=[.21,.64]). Listeners also perceived unresponsive 'ghost partners' to lead live musicians in *one-way* duos (M=.14, SD = 0.03, 95% CI = [0.08, 0.20]), whereas leadership was perceived to be more evenly distributed in coupled duos (effect of condition on deviation of leadership ratings from neutral: M = .14, SD = .03, 95% CI = [.08, .19]). However, listeners did not guess the correct condition above chance level (M = .03, SD = 0.09, 95% CI = [-0.14, 0.21]). These results held equally for both populations of listeners, as no effects of subject type were observed.

Mutual Coupling Promotes Synchrony

[2] Henceforth "duo" refers to a pair of performers and "duet" refers to a particular piece produced by a duo. Each coupled duo yoked two one-way duos, same for duets.

How does coupling influence musicians' ability to synchronize with one another? Asynchronies between "near-simultaneous" onsets (co-occurring within 100 milliseconds) played by co-performers were measured throughout all duets in each condition. Near-zero asynchronies indicate close temporal alignment, while asynchronies of larger magnitude reflect less precise synchronization. As depicted in Figure 1, asynchronies in *coupled* trials are peaked around zero (red distribution), whereas asynchronies in *one-way* trials are more widely distributed throughout the +/- 100 ms range (blue distribution) (KS.test D = 0.024, p-value <.01), indicating that mutually coupled musicians achieved more precise synchronization compared to musicians in the overdubbed condition. We were also curious about leader-follower asymmetries in one-way duos, as previous studies have reported that supporting musicians lag behind lead musicians in certain composed musical contexts Keller and Appel (2010). However no such effects were observed here; the distribution of asynchronies in *one-way* duets was not significantly asymmetric around 0 in one direction or the other.

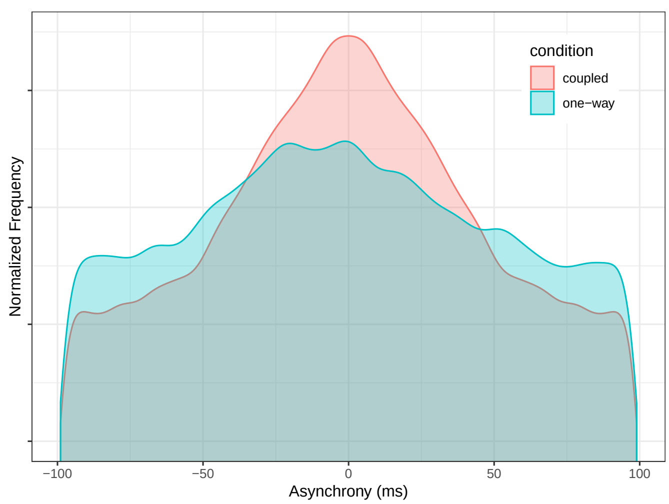

Figure 1. Mutual coupling facilitates precise synchronization. Distribution of asynchronies between co-performers' near-simultaneous (within 100 ms) note onsets throughout all trials in each condition. Asynchronies are more tightly clustered around 0 in coupled trials, indicating more precise temporal alignment.

Activity Matching

Lagged cross-correlation of co-performers' onset density was computed to analyze how musicians responded to one another's rhythmic activity level. Onset density contributes to the perception of musical tension Farbood (2012). A frenzied musical passage comprising many notes in rapid succession would yield high onset density, whereas a more sparse, mellow passage would yield low onset density. Onset density time series were computed for each individual note sequence using a 2-second sliding window, with a 0.2-second hop size. Figure 2 depicts lagged cross-correlations, averaged across all duets in each condition. Cross-correlation was positive overall (cross-correlation averaged across +/-20 second lag range: M = .39, SD = .04, 95 % CI = [.31,.47]), but significantly higher in *coupled* duos (red curve) (fixed effect of condition: M = -.13, SD = .04, CI = [-.21,-.06]). These results indicate a general tendency for musicians to match the onset density of their partners, which was exaggerated in mutually coupled duos.

Within one-way duos, cross-correlation was significantly higher at positive, 'ghost-to-live' lags (onset density ghost recording correlated with future onset density of live musician) compared to negative, 'live-to-ghost' lags (effect of direction: M = .05, SD = .01, 95% CI = [.02,.08]). This reflects the underlying asymmetry in *one-way* duets: live musicians were responsive to notes of ghost recordings but not the other way around. As reported in the Supporting Information, a complementary Granger Causality analysis also revealed greater ghost-to-live versus live-to-ghost Granger causality in *one-way* duos. Lastly, Figure 2 reveals a dip in cross-correlation for coupled duets at simultaneous timepoints, but this was not statistically significant.

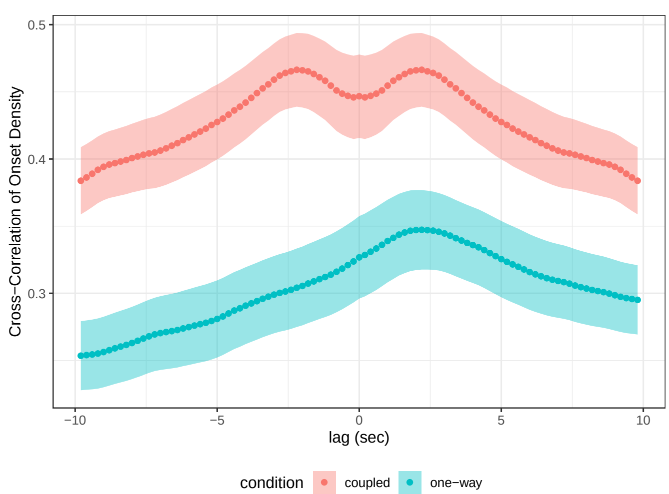

Figure 2. Musicians match the activity level of their partners. Points represent mean lagged cross-correlation across all trials within each condition. Error ribbons denote standard error of the mean. Positive lags in *one-way* trials represent the correlation of ghost recording onset density with future onset density of live musicians (ghost-to-live) and vice versa for negative lags (live-to-ghost).

Emergence and Directed Flow of Tonal Information

A previously established model of tonal structure (see Methods and SI) was adapted to provide a measure of *tonal consonance*, quantifying how collections of notes sound on a continuum from unstable/dissonant to stable/consonant Chew et al. (2014); Herremans et al. (2016). Time series of Combined Consonance (consonance of merged music streams from both players in a duo) were computed with a sliding window[3] Emergent Consonance (EC) was operationalized as Combined Consonance minus average consonance of each individual music stream. EC captures the consonance arising from the *interaction* of pitches played by collaborating musicians. A situation in which each pianist plays self-consonant notes that clash with one another would result in low EC (e.g. {C,E,G} and {F#,A#,C#} are consonant on their own but {C,E,G,F#,A#,C#} is highly dissonant), whereas a situation in which each pianist plays dissonant notes that stabilize one another when sounded together would result in high EC (e.g. {C,B} and {E,G} have low average consonance but {C,E,G,B} has high consonance because it is tonicized to a Cmaj7 chord). Negative EC values indicate that Combined Consonance is less consonant than the average Individual Consonance and can be interpreted as emergent *dissonance*. Less negative values can be interpreted as indexing greater EC (less emergent dissonance) compared to more negative values.

A novel lagged consonance analysis was used to quantify how musicians harmonized with one another's notes as a function of interaction condition. Lagged consonance was computed by shifting individual note sequences of co-performers relative to one another, computing Combined and Emergent Consonance time series of the merged pitch collections with a sliding window, and then averaging over time to get a single consonance value per piece at each lag (5 second sliding window and 2 second hop size were used, although these results were robust across a range of window sizes, as documented in the SI). This analysis captures the directed flow of tonal information, as it quantifies the degree to which individuals harmonized with the preceding notes of their partner. For example, Player A might harmonize with Player B's past notes but not the other way around, which would be reflected in high consonance for B-to-A lags but not A-to-B lags. Lagged consonance was computed for every trial in each condition with lags in the range of +/-20 seconds, spaced by increments of 2 seconds. Positive lags in one-way duos

[3] A range of window sizes (2, 5 and 10 seconds) were evaluated, with a hop size of 2 seconds. Many of the following reported results were robust across all window sizes, but we indicate cases where this isn't true.

correspond to evaluating past notes of the ghost recording with future notes of the live musician (ghost-to-live) and vice versa for negative lags (live-to-ghost). The beginnings and endings of pieces (first and last 10%) were discarded to avoid boundary effects.

Figure 3 depicts average lagged Emergent Consonance (EC) by condition. Time lags are plotted on the x-axis, and the y-axis represents average lagged Emergent Consonance throughout all duets in each condition. EC is essentially symmetric around 0 seconds (simultaneous playing) for *coupled* trials (red curve), but significantly higher in ghost-to-live (positive) lags compared to live-to-ghost (negative) lags for *one-way* trials (blue curve) (effect of lag sign on EC averaged across negative and positive lags: M = -2.90e-3, SD = 1.18e-3, 95% CI = [-5.22e-3, -5.35e-3]). This asymmetry in *one-way* trials was also found with respect to Combined Consonance (effect of lag sign on average CC: M = -3.45e-3, SD = 1.45e-3, 95% CI = [-6.32e-3, -6.19e-4]). These results reflect the underlying causal entanglements of each condition. Live musicians in *one-way* trials responded to ghost recordings by harmonizing with their past notes, but ghost recordings could not respond to notes of live musicians. There was no such asymmetry in coupled trials, because musicians were mutually responsive. As suggested by the difference in height between red (coupled) and blue (one-way) data points in Figure 3, EC was significantly higher overall in coupled versus one-way duos (effect of condition on simultaneous EC: M = -1.09e-2, SD = 4.91e-3, 95% CI = [-2.05e-2, -1.01e-3]), although this effect was not significant with respect to Combined Consonance (M = -1.51e-2, SD = 9.50e-3, 95% CI = [-3.41e-2, 3.21e-3]). In sum, coupled improvisers mutually harmonized with one another's preceding notes, and this dynamic supported more consonant harmonization between them.

DISCUSSION

This study examined how music produced by collaboratively improvising musicians is shaped by underlying patterns of coordination. Professional jazz pianists improvised in two duo conditions: a *coupled* condition in which they improvised together simultaneously, and an "overdubbed" (*one-way*) condition which precluded mutual adaptation because improvisers were recorded sequentially. Our analyses show that *coupled* duos achieved greater alignment of their note onsets and more consonant tonal coordination. These results were paralleled in the subjective experience of the performers and naive listeners, who preferred *coupled* duets despite being blind to condition.

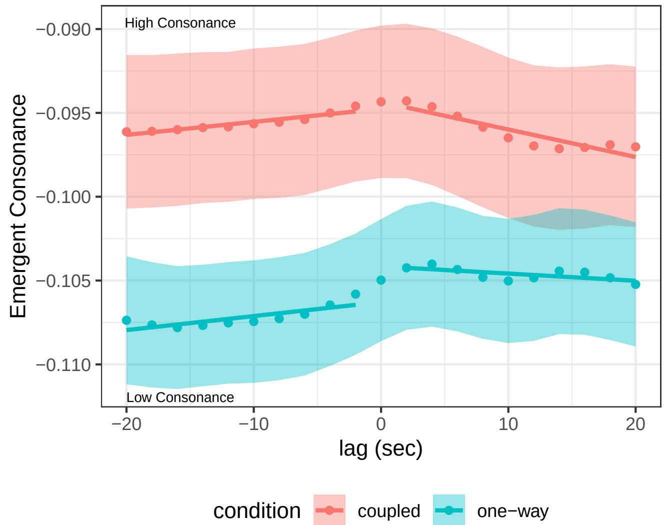

Figure 3. Lagged consonance analysis reveals musicians harmonize with preceding notes of their partner. Negative lags correspond to notes of the live musician merged with future notes of the ghost recording (live-to-ghost) and vise versa for positive lags (ghost-to-live). Points denote average EC at a given lag across every piece within each condition, error bars denote standard error of the mean. Linear fits of EC by lag are shown for negative and positive lags in each condition.

Performers and listeners demonstrated systematic insight into the different causal entanglements of each condition. Leadership was rated as evenly distributed amongst *coupled* duos, but listeners perceived “ghost partners” as leading live musicians and performers rated themselves as playing more of a follower/supporter role in one-way duets. These listener results are remarkable in light of the fact that they were unable to guess which condition music samples were produced in above chance-level. Listener perception was thus implicitly influenced by the presence or absence of mutual coupling, without their conscious awareness.

Coupled duos synchronized their note onsets more precisely than one-way duos, as in previous studies which showed that bidirectional coordination promotes synchronization in finger-tapping tasks Konvalinka et al. (2010) and scored music performance Demos et al. (2017); Goebl and Palmer (2009). Here this phenomenon is observed in freely improvising musicians, with no explicit synchronization objective. Rather, precise synchronization emerged spontaneously, in service of the higher-level goal of collectively generating compelling music. Previous findings have also suggested that humans have an innate predisposition to entrain rhythms in social contexts Kirschner and Tomasello (2009), which could elucidate our result insofar as pianists may have sensed a lack of live responsiveness in their partners in *one-way* duets.

Mutual coupling supported note onset alignment at longer timescales as well. A cross-correlation analysis of onset density revealed that improvisers tended to match the rhythmic activity of their partners, and this tendency was significantly stronger in coupled duos. This relates to findings in non-musical JA domains. Previous dyadic conversation studies have shown that people spontaneously entrain their movements, and mimic one another's facial expressions, manual gestures, eye gaze and acoustic speech characteristics when verbally interacting with one another Abney et al. (2014); Louwerse et al. (2012); D. C. Richardson and Dale (2005); Shockley et al. (2009, 2003). Behavioral alignment has been proposed to foster successful interaction by signaling affiliative attitudes Demos et al. (2012); Hove and Risen (2009), and offloading predictive emulation (i.e. of a conversation partner's future behavior) onto one's own behavior Garrod and Pickering (2009); the temporal alignment observed here may serve these same interpersonal functions in improvised musical interactions.

Our onset density cross-correlation analysis also inferred different profiles of directional influence for each interaction condition. Cross-correlation was symmetric between coupled partners, but there was an asymmetry in one-way duos such that onset density of the live musician correlated with past onset density of the ghost partner (prerecorded track) but not vice versa. This result adds to previous demonstrations that causal influence in performing music ensembles is reflected in the movements and music of co-performers. This has been shown numerous times in the context of composed music Chang et al. (2017); Demos et al. (2017); Keller and Appel (2010), and the work of Aucouturier and Canonne (2017) suggested that leader-follower roles induced by experimentally manipulated social attitudes (e.g. caring, dominant) are reflected in sound envelopes (loudness) of improvising musicians. However, this latter finding was somewhat speculative because inter-musician coupling was not explicitly manipulated. In contrast, our overdubbed interaction condition provides a ground-truth to verify our analysis against.

Analogous findings were uncovered in the realm of abstract tonal structure. A novel lagged consonance analysis demonstrated that musicians harmonized with the past notes of their partners. This occurred mutually in *coupled* duos but asymmetrically in *one-way* duos, where live musicians harmonized with preceding notes of the ghost recording, but not vice versa. Causal influence between improvisers was thus reflected not just in their rhythms, but also in the notes they played and the directed exchange of tonal information. Additionally, simultaneous Emergent Consonance was significantly greater in *coupled* duos, suggesting that the ability to mutually adapt to one another's previous notes promoted robust tonal coordination.

Importantly, our consonance analysis detected not just alignment, but complementary tonal coordination as well. Consonance is not only achieved when musicians play the same pitch, but also when they play complementary sets of pitches that combine to produce consonant harmonies. The tonal coupling observed here can be understood in terms of interpersonal synergies, which have been proposed to emerge in interacting groups whose individuals co-constrain one another in support of group-level objectives Hasson and Frith (2016); Riley, Richardson, Shockley, and Ramenzoni (2011). In this case, note selection is co-constrained between collaboratively improvising musicians in order to generate tonal structure. Our consonance analysis contributes an important extension to previous analyses of naturalistic JA, which have primarily operationalized coordination in terms of behavioral matching, using techniques like cross-correlation and cross recurrence analysis Dale and Spivey (2006); Louwerse et al. (2012); Paxton and Dale (2013); D. C. Richardson and Dale (2005); D. C. Richardson et al. (2007). Here we demonstrate the feasability of using domain-specific measures (i.e. a tonal consonance model informed by music theory) to assess complementary coordination in support of abstract, functional properties at the group-level (i.e. emergent tonal structure). While there can be no doubt that alignment is an important interpersonal mechanism, more work of this kind is needed to investigate complementary coordination in naturalistic JA contexts Hasson and Frith (2016).

Successful coordination is difficult to operationalize in freely improvised music, because it is not explicitly clear what the intentions of musicians are. We analyzed rhythmic alignment and tonal consonance because they are basic musical elements, and we were able to operationalize them while imposing minimal musical assumptions (atonal music would be rated low consonance, onset density works for pulsed and non-pulsed music). The goal of participants was to generate compelling music, as they would strive for in a typical performance, but they were not explicitly instructed to synchronize note onsets or produce consonant harmonies. In fact, some level of musical tension and dissonance is typically desired. This being said, we observed robust effects that mutual coupling promoted temporal alignment and emergent tonal consonance overall. We also observed directional effects on these features consistent with the ground-truth unidirectional influence from recording to musician in *one-way* duets.

Furthermore, these results were paralleled in the subjective experience of professional improvisers and naive listeners with no particular background in jazz music, who preferred mutually coupled duos, and correctly inferred leadership roles in both conditions.

Taken together, these results suggest that coupled dyads achieved enhanced, bidirectional temporal and tonal coordination, which supported the higher-level goal of generating compelling music. This extends previous investigations of mutual coupling in idealized experimental paradigms, such as finger tapping Konvalinka et al. (2010) and the improvised mirror game Noy et al. (2011), into the rich, naturalistic setting of unconstrained musical improvisation. More specifically, our findings directly implicate the common studio recording technique of overdubbing – which we show results in systematically different music than live, coupled interaction. Lastly, our measures of expert musical coordination can be incorporated into the design of generative AI music systems to make them more human-like and more musical Datseris et al. (2019); Gillick et al. (2019); Hawthorne et al. (2019); Hennig (2014); Hennig et al. (2011); Huang et al. (2019); A. Roberts et al. (2019).

REFERENCES

Abney, D. H., Paxton, A., Dale, R., & Kello, C. T. (2014). Complexity matching in dyadic conversation. *Journal of Experimental Psychology: General*, *143*(6), 2304.

Andrade-Lotero, E., & Goldstone, R. L. (n.d.). Self-organized division of cognitive labor.

Aucouturier, J.-J., & Canonne, C. (2017). Musical friends and foes: The social cognition of affiliation and control in improvised interactions. *Cognition*, *161*, 94–108.

Barnett, L., & Seth, A. K. (2014). The mvgc multivariate granger causality toolbox: a new approach to granger-causal inference. *Journal of neuroscience methods*, *223*, 50–68.

Berkowitz, A. (2010). *The improvising mind: Cognition and creativity in the musical moment*. Oxford University Press.

Borgo, D. (2005). *Sync or swarm: Improvising music in a complex age*. A&C Black.

Bürkner, P.-C., et al. (2017). brms: An r package for bayesian multilevel models using stan. *Journal of statistical software*, *80*(1), 1–28.

Canonne, C., & Aucouturier, J.-J. (n.d.). Play together, think alike: Shared mental models in expert music improvisers. , *44*(3), 544–558.

Chang, A., Livingstone, S. R., Bosnyak, D. J., & Trainor, L. J. (n.d.). Body sway reflects leadership in joint music performance. , *114*(21), E4134–E4141.

Chang, A., Livingstone, S. R., Bosnyak, D. J., & Trainor, L. J. (2017). Body sway reflects leadership in joint music performance. *Proceedings of the National Academy of Sciences*, 201617657.

Chew, E. (2005). Regards on two regards by messiaen: post-tonal music segmentation using pitch context distances in the spiral array. *Journal of New Music Research*, *34*(4), 341–354.

Chew, E., et al. (2014). Mathematical and computational modeling of tonality. *AMC*, *10*, 12.

Council, N. R., et al. (2015). *Enhancing the effectiveness of team science*. National Academies Press.

Curioni, A., Vesper, C., Knoblich, G., & Sebanz, N. (2019). Reciprocal information flow and role distribution support joint action coordination. *Cognition*, *187*, 21–31.

Dale, R., & Spivey, M. J. (2006). Unraveling the dyad: Using recurrence analysis to explore patterns of syntactic coordination between children and caregivers in conversation. *Language Learning*, *56*(3), 391–430.

Datseris, G., Ziereis, A., Albrecht, T., Hagmayer, Y., Priesemann, V., & Geisel, T. (2019). Microtiming deviations and swing feel in jazz. *Scientific Reports*, *9*(1), 1–10.

D'Ausilio, A., Badino, L., Li, Y., Tokay, S., Craighero, L., Canto, R., … Fadiga, L. (2012). Leadership in orchestra emerges from the causal relationships of movement kinematics. *PLoS one*, *7*(5), e35757.

Demos, A. P., Carter, D. J., Wanderley, M. M., & Palmer, C. (2017). The unresponsive partner: roles of social status, auditory feedback, and animacy in coordination of joint music performance. *Frontiers in psychology*, *8*, 149.

Demos, A. P., Chaffin, R., Begosh, K. T., Daniels, J. R., & Marsh, K. L. (2012). Rocking to the beat: Effects of music and partner's movements on spontaneous interpersonal coordination. *Journal of Experimental Psychology: General*, *141*(1), 49.

Eerola, T., Jakubowski, K., Moran, N., Keller, P. E., & Clayton, M. (n.d.). Shared periodic performer movements coordinate interactions in duo improvisations. , *5*(2), 171520.

Eerola, T., Jakubowski, K., Moran, N., Keller, P. E., & Clayton, M. (2018). Shared periodic performer movements coordinate interactions in duo improvisations. *Royal Society open science*, *5*(2), 171520.

Farbood, M. M. (2012). A parametric, temporal model of musical tension. *Music Perception: An Interdisciplinary Journal*, *29*(4), 387–428.

Fusaroli, R., Raczaszek-Leonardi, J., & Tylén, K. (n.d.). Dialog as interpersonal synergy. , *32*, 147–157.

Fusaroli, R., & Tylén, K. (2016). Investigating conversational dynamics: Interactive alignment, interpersonal synergy, and collective task performance. *Cognitive science*, *40*(1), 145–171.

Fusaroli, R., & Tylén, K. (n.d.). Investigating conversational dynamics: Interactive alignment, interpersonal synergy, and collective task performance. , *40*(1), 145–171.

Garrod, S., & Pickering, M. J. (n.d.). Joint action, interactive alignment, and dialog. , *1*(2), 292–304.

Garrod, S., & Pickering, M. J. (2009). Joint action, interactive alignment, and dialog. *Topics in Cognitive Science*, *1*(2), 292–304.

Gillick, J., Roberts, A., Engel, J., Eck, D., & Bamman, D. (2019). Learning to groove with inverse sequence transformations. *arXiv preprint arXiv:1905.06118*.

Goebl, W., & Palmer, C. (n.d.). Synchronization of timing and motion among performing musicians. , *26*(5).

Goebl, W., & Palmer, C. (2009). Synchronization of timing and motion among performing musicians. *Music Perception: An Interdisciplinary Journal*, *26*(5), 427–438.

Goldstone, R. L., & Gureckis, T. M. (2009). Collective behavior. *Topics in cognitive science*, *1*(3), 412–438.

Guimera, R., Uzzi, B., Spiro, J., & Amaral, L. A. N. (2005). Team assembly mechanisms determine collaboration network structure and team performance. *Science*, *308*(5722), 697–702.

Hasson, U., & Frith, C. D. (2016). Mirroring and beyond: coupled dynamics as a generalized framework for modelling social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1693), 20150366.

Hasson, U., Ghazanfar, A. A., Galantucci, B., Garrod, S., & Keysers, C. (2012). Brain-to-brain coupling: a mechanism for creating and sharing a social world. *Trends in cognitive sciences*, *16*(2), 114–121.

Hatano, G., & Oura, Y. (2003). Commentary: Reconceptualizing school learning using insight from expertise research. *Educational researcher*, *32*(8), 26–29.

Hawthorne, C., Stasyuk, A., Roberts, A., Simon, I., Huang, A., Dieleman, S., … Eck, D. (2019). Enabling factorized piano music modeling and generation with the maestro dataset.. Retrieved from https://arxiv.org/pdf/1810.12247

Hennig, H. (2014). Synchronization in human musical rhythms and mutually interacting complex systems. *Proceedings of the National Academy of Sciences*, *111*(36), 12974–12979.

Hennig, H., Fleischmann, R., Fredebohm, A., Hagmayer, Y., Nagler, J., Witt, A., … Geisel, T. (2011). The nature and perception of fluctuations in human musical rhythms. *PloS one*, *6*(10), e26457.

Herremans, D., Chew, E., et al. (2016). Tension ribbons: Quantifying and visualising tonal tension.

Heylighen, F. (2013). Self-organization in communicating groups: the emergence of coordination, shared references and collective intelligence. In *Complexity perspectives on language, communication and society* (pp. 117–149). Springer.

Hove, M. J., & Risen, J. L. (n.d.). It's all in the timing: Interpersonal synchrony increases affiliation. , *27*(6), 949–960.

Hove, M. J., & Risen, J. L. (2009). It's all in the timing: Interpersonal synchrony increases affiliation. *Social cognition*, *27*(6), 949–960.

Huang, C.-Z. A., Vaswani, A., Uszkoreit, J., Shazeer, N., Simon, I., Hawthorne, C., … Eck, D. (2019). Music transformer: Generating music with long-term structure.. Retrieved from https://arxiv.org/abs/1809.04281

Johnson-Laird, P. N. (1988). Freedom and constraint in creativity. *The nature of creativity: Contemporary psychological perspectives*, *202*.

Johnson-Laird, P. N. (2002). How jazz musicians improvise. *Music Perception: An Interdisciplinary Journal*, *19*(3), 415–442.

Johnson-Laird, P. N., Kang, O. E., & Leong, Y. C. (2012). On musical dissonance. *Music Perception: An Interdisciplinary Journal*, *30*(1), 19–35.

Ju, W., & Leifer, L. (2008). The design of implicit interactions: Making interactive systems less obnoxious. *Design Issues*, *24*(3), 72–84.

Keller, P. E. (2014). Ensemble performance: Interpersonal alignment of musical expression. *Expressiveness in music performance: Empirical approaches across styles and cultures*, 260–282.

Keller, P. E., & Appel, M. (2010). Individual differences, auditory imagery, and the coordination of body movements and sounds in musical ensembles. *Music Perception: An Interdisciplinary Journal*, *28*(1), 27–46.

Kirschner, S., & Tomasello, M. (2009). Joint drumming: social context facilitates synchronization in preschool children. *Journal of experimental child psychology*, *102*(3), 299–314.

Knoblich, G., Butterfill, S., & Sebanz, N. (2011). Psychological research on joint action: theory and data. In *Psychology of learning and motivation* (Vol. 54, pp. 59–101). Elsevier.

Konvalinka, I., Vuust, P., Roepstorff, A., & Frith, C. D. (2010). Follow you, follow me: continuous mutual prediction and adaptation in joint tapping. *Quarterly journal of experimental psychology*, *63*(11), 2220–2230.

Limb, C. J., & Braun, A. R. (2008). Neural substrates of spontaneous musical performance: An fmri study of jazz improvisation. *PLoS one*, *3*(2).

Linson, A., Dobbyn, C., Lewis, G. E., & Laney, R. (2015). A subsumption agent for collaborative free improvisation. *Computer Music Journal*, *39*(4), 96–115.

Louwerse, M. M., Dale, R., Bard, E. G., & Jeuniaux, P. (2012). Behavior matching in multimodal communication is synchronized. *Cognitive science*, *36*(8), 1404–1426.

Marwan, N., Romano, M. C., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics reports*, *438*(5-6), 237–329.

Noy, L., Dekel, E., & Alon, U. (2011). The mirror game as a paradigm for studying the dynamics of two people improvising motion together. *Proceedings of the National Academy of Sciences*, *108*(52), 20947–20952.

Palmer, C., & Zamm, A. (2017). Empirical and mathematical accounts. *The Routledge Companion to Embodied Music Interaction*, 370.

Paulus, P. B., Levine, D. S., Brown, V., Minai, A. A., & Doboli, S. (2010). Modeling ideational creativity in groups: Connecting cognitive, neural, and computational approaches. *Small Group Research*, *41*(6), 688–724.

Paxton, A., & Dale, R. (2013). *Argument disrupts interpersonal synchrony.* SAGE Publications Sage UK: London, England.

Pearce, M., & Rohrmeier, M. (2012). Music cognition and the cognitive sciences. *Topics in cognitive science*, *4*(4), 468–484.

Pickering, M. J., & Garrod, S. (n.d.). An integrated theory of language production and comprehension. , *36*(4), 329–347.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and brain sciences*, *27*(2), 169–190.

Pressing, J. (1988). Improvisation: methods and models. *John A. Sloboda (Hg.): Generative processes in music, Oxford*, 129–178.

Rasch, R. A. (n.d.). *Synchronization in performed ensemble music*.

Rasch, R. A. (1979). Synchronization in performed ensemble music. *Acta Acustica united with Acustica*, *43*(2), 121–131.

Rebsamen, B., Guan, C., Zhang, H., Wang, C., Teo, C., Ang, M. H., & Burdet, E. (2010). A brain controlled wheelchair to navigate in familiar environments. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *18*(6), 590–598.

Repp, B. H., & Su, Y.-H. (n.d.). Sensorimotor synchronization: A review of recent research (2006–2012). , *20*(3), 403–452.

Richardson, D. C., & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive science*, *29*(6), 1045–1060.

Richardson, D. C., Dale, R., & Kirkham, N. Z. (2007). The art of conversation is coordination. *Psychological science*, *18*(5), 407–413.

Richardson, D. C., Dale, R., & Tomlinson, J. M. (n.d.). Conversation, gaze coordination, and beliefs about visual context. , *33*(8), 1468–1482.

Richardson, M. J., Dale, R., & Marsh, K. L. (2014). Complex dynamical systems in social and personality psychology. *Handbook of research methods in social and personality psychology*, *253*.

Richardson, M. J., Harrison, S. J., Kallen, R. W., Walton, A., Eiler, B. A., Saltzman, E., & Schmidt, R. (2015). Self-organized complementary joint action: Behavioral dynamics of an interpersonal collision-avoidance task. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(3), 665.

Richardson, M. J., Marsh, K. L., Isenhower, R. W., Goodman, J. R., & Schmidt, R. C. (2007). Rocking together: Dynamics of intentional and unintentional interpersonal coordination. *Human movement science*, *26*(6), 867–891.

Richardson, M. J., Marsh, K. L., Isenhower, R. W., Goodman, J. R. L., & Schmidt, R. C. (n.d.). Rocking together: Dynamics of intentional and unintentional interpersonal coordination. , *26*(6), 867–891.

Riley, M. A., Richardson, M., Shockley, K., & Ramenzoni, V. C. (2011). Interpersonal synergies. *Frontiers in psychology*, *2*, 38.

Roberts, A., Huang, A., Hawthorne, C., Howcroft, J., Wexler, J., Hong, L., & Dinculescu, M. (2019). Approachable music composition with machine learning at scale. In *Proceedings of the 20th international society for music information retrieval conference (ismir).*

Roberts, M. E., & Goldstone, R. L. (2011). Adaptive group coordination and role differentiation. *PLoS One*, *6*(7).

Schmidt, R. C., & Richardson, M. J. (2008). Dynamics of interpersonal coordination. In *Coordination: Neural, behavioral and social dynamics* (pp. 281–308). Springer.

Sebanz, N., Bekkering, H., & Knoblich, G. (n.d.). Joint action: bodies and minds moving together. , *10*(2), 70–76.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in cognitive sciences*, *10*(2), 70–76.

Sebanz, N., & Knoblich, G. (n.d.). Prediction in joint action: What, when, and where. , *1*(2), 353–367.

Sebanz, N., & Knoblich, G. (2009). Prediction in joint action: What, when, and where. *Topics in Cognitive Science*, *1*(2), 353–367.

Shockley, K., Richardson, D. C., & Dale, R. (2009). Conversation and coordinative structures. *Topics in Cognitive Science*, *1*(2), 305–319.

Shockley, K., Santana, M.-V., & Fowler, C. A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(2), 326.

Valdesolo, P., Ouyang, J., & DeSteno, D. (2010). The rhythm of joint action: Synchrony promotes cooperative ability. *Journal of Experimental Social Psychology*, *46*(4), 693–695.

Walton, A. E., Richardson, M. J., Langland-Hassan, P., & Chemero, A. (2015). Improvisation and the self-organization of multiple musical bodies. *Frontiers in psychology*, *6*, 313.

Walton, A. E., Washburn, A., Langland-Hassan, P., Chemero, A., Kloos, H., & Richardson, M. J. (2018). Creating time: Social collaboration in music improvisation. *Topics in cognitive science*, *10*(1), 95–119.