Charles Patrick Martin*
and Jim Torresen†

*Research School of Computer Science
The Australian National University
Canberra, ACT, 2601, Australia
†Department of Informatics
University of Oslo
Postboks 1080, Blindern, 0316 Oslo, Norway
charles.martin@anu.edu.au,
jimtoer@ifi.uio.no

Data-Driven Analysis of Tiny Touchscreen Performance with MicroJam

Abstract: The widespread adoption of mobile devices, such as smartphones and tablets, has made touchscreens a common interface for musical performance. Although new mobile music instruments have been investigated from design and user experience perspectives, there has been little examination of the performers' musical output. In this work, we introduce a constrained touchscreen performance app, MicroJam, designed to enable collaboration between performers, and engage in a data-driven analysis of more than 1,600 performances using the app. MicroJam constrains performances to five seconds, and emphasizes frequent and casual music-making through a social media–inspired interface. Performers collaborate by replying to performances, adding new musical layers that are played back at the same time. Our analysis shows that users tend to focus on the center and diagonals of the touchscreen area, and that they tend to swirl or swipe rather than tap. We also observe that, whereas long swipes dominate the visual appearance of performances, the majority of interactions are short with limited expressive possibilities. Our findings enhance our understanding of how users perform in touchscreen apps and could be applied in future app designs for social musical interaction.

Popular social media apps for mobile devices have allowed millions of users to engage with creative production of images and text. These devices' cameras, touchscreens, powerful processors, and portability suggest on-the-go creativity, and it would appear that straightforward sharing with friends, or a wider network of followers, is a key factor in encouraging users to create content of all forms. Given the many affordances of mobile devices, they are recognized as suitable platforms for making music (Tanaka 2010). Many creative mobile music apps have appeared in recent years, for example, Magic Fiddle (Wang, Oh, and Lieber 2011), Ocarina (Wang 2014), Orphion (Trump and Bullock 2014), PhaseRings (Martin et al. 2016), and TC-11 (Schlei 2012). Despite interest in these apps, mobile musical creation has not yet been adopted as an integrated element of mainstream social media, although some music apps have provided social features. Furthermore, few musical apps have attempted to emphasize ensemble, rather than individual, performance, even though group music-making is often seen as a valuable social activity.

In this article, we present the design for MicroJam, a collaborative and social mobile music-making app, and an analysis of more than 1,600 touchscreen performances that have been created so far. The design of MicroJam emphasizes casual, frequent, and social performance. As shown in Figure 1, the app features a simple touchscreen interface for making electronic music where skill is not a necessary prerequisite for interaction. (Color images of all figures are available at https://www.mitpressjournals.org/doi/suppl/COMJ_a_00536-Martin.) MicroJam departs from other touchscreen instruments by imposing limits on the musical compositions that are possible; most importantly, performances are limited to five seconds in length. These "tiny" performances are uploaded automatically to encourage improvisation and creation rather than editing. Users can reply to others' performances by recording a new layer, combining social interaction with ensemble music-making.

In the sections that follow, we motivate MicroJam's design with a discussion of music-making in social media, and the possibilities for asynchronous and distributed collaborations with mobile musical interfaces. We describe the app's design and the interactive music mappings of the nine synthesized instruments that are made available in its interface.

Figure 1. MicroJam allows
users to create short
touchscreen performances
that are shared with other
users. Replies to
performances form
distributed and
asynchronous duets. A
video demo of MicroJam is
available, together with
color versions of all figures,
at https://www.mitpress
journals.org/doi/suppl
/COMJ_a_00536-Martin.

We also formalize the concept of tiny touchscreen performances. The main contribution of this work is in our analysis of a data set of more than 1,600 tiny performances saved on the app's cloud database by testers and users. We consider this data set from several levels of abstraction: individual touchscreen interaction events, aggregated touchscreen gestures, and whole performances. This analysis allows us to draw conclusions about how these users perform in MicroJam, and to characterize the musical behavior in the tiny performance format. Although our app design draws on themes introduced by other authors, this is the first time that a systematic and data-driven analysis of a large data set of touchscreen performances has been published for such a system. Our findings, as well as our data-driven method, could be applied to the design and study of other touchscreen music apps and other types of interactive music systems, as well as to future revisions of MicroJam.

This article is a revised and extended version of a previous conference paper (Martin and Torresen 2017). In this new research, we present the fully developed app and an analysis of touchscreen performances.

Background

Commodity mobile devices such as smartphones and tablets have often been reframed as musical instruments for research, artistic exploration, and entertainment, forming the field of mobile music (Gaye et al. 2006). The sensors and multitouch screens of smartphones provide many affordances for new kinds of musical software (Essl and Rohs 2009; Tanaka 2010) and the ubiquity of these devices increases the possibility of musical participation by a wide audience (Essl and Lee 2018). Most mobile music apps have used the touchscreen as an expressive musical controller. Some apps, such as Magic Fiddle (Wang, Oh, and Lieber 2011), imitate an existing musical instrument, but others, such as Crackle (Reus 2011) or TC-11 (Schlei 2012), have defined new ways to connect interaction on the touchscreen to sound synthesis algorithms.

Although the design and evaluation of mobile music apps have been reported in the academic literature for surveys (see, e.g., Gaye et al. 2006; John 2013; Essl and Lee 2018), few analyses exist specifically of the music created with these systems. Evaluations of these systems tend to focus on either the design and gesture-to-sound mappings (e.g., d'Alessandro et al. 2012), or on the experience of the musicians using the apps (e.g., Martin et al. 2016). It has previously been argued that archives of touch-screen control data can go beyond audio in terms of analysis of tablet musical performances (Martin and Gardner 2016). Jeff Smith (2013) analyzed a database of 800,000 MIDI performances of popular songs played in the Magic Piano tablet app and identified musical behaviors based on geographical location of the performer. Similar data from motion capture systems have previously been used as a basis for analysis of human interactions between movement and sound (Kelkar and Jensenius 2018). In this work, we perform an analysis on mobile app touchscreen data to understand how users interact musically with MicroJam.

Social Music-Making and Constraints

Many social media platforms emphasize the value of constrained contributions by users. Twitter famously limited written notes to 140 characters (Gligorić, Anderson, and West 2018), Instagram constrained images to square format, and the now-defunct Vine platform only allowed six-second microvideos (Redi et al. 2014). Constraints are often thought to lead to increased variability and

creativity in the arts (Stokes 2008) and in electronic musical instruments (Gurevich, Marquez-Borbon, and Stapleton 2012). Posts in Twitter or Instagram are intended to be frequent, casual, and ephemeral, and it could be that these constrained formats have helped these apps to attract millions of users and encourage their creativity in the written word or photography. And whereas social media is often used to promote music (Dewan and Ramaprasad 2014), the actual making of music has yet to become an important creative part of the social-media landscape outside of a number of standalone apps.

Although music is often seen as an activity where accomplishment takes practice and concerted effort, casual musical experiences are known to be valuable. Accessible music-making, such as percussion drum circles, can be used for music therapy (Scheffel and Matney 2014). Augmented-reality instruments (Correa et al. 2009) and touchscreen instruments (Favilla and Pedell 2013) have also been used for therapeutic and casual music-making. Apps such as Ocarina (Wang 2014) and Pyxis Minor (Barraclough, Carnegie, and Kapur 2015) have shown that simple touchscreen interfaces can be successful for exploration by novice users as well as supporting sophisticated expressions by practiced performers.

Collaboration in Mobile Music

Telecommunication networks have long been used to facilitate and democratize musical collaboration (Kim-Boyle 2009); for example, Max Neuhaus's *Radio Net* (http://www.ubu.com/sound/neuhaus _radio.html) used the sounds of radio listeners collectively calling in and whistling into their phone. Later, computer networks enabled software instruments to be distributed to collaborating performers over the Web, such as in Sergi Jordà's FMOL system (Jordà 1999). Musical apps have taken advantage of mobile devices' always-on Internet to include various kinds of collaboration. Smule's Leaf Trombone app introduced a "world stage" (Wang et al. 2015), where users from around the world were invited to critique performances with emoticons and short text comments. Mobile

devices are also used in ensemble situations such as the Stanford Mobile Phone Orchestra (MoPho, see Oh et al. 2010), Viscotheque (Swift 2013), Pocket Gamelan (Schiemer and Havryliv 2007), ChoirMob (d'Alessandro et al. 2012), and Ensemble Metatone (Martin, Gardner, and Swift 2015). In these examples, however, the musicians played together in a standard concert situation and are not geographically dispersed.

Given that mobile devices are often carried by users at all times, it is natural to ask whether mobile ensemble experiences can be achieved outside of a typical concert. Collaborative configurations can be framed in a time–space matrix, that is, occurring together or apart in both location and time (Greenberg and Roseman 1998). Although networked music performances (Carôt, Rebelo, and Renaud 2007), with performers distributed in space, are well documented (e.g., Bryan-Kinns 2004), collaborations across time are less thoroughly explored. Some location-based performances (Behrendt 2012) have allowed asynchronous collaboration—for instance, in AuRal (Allison and Dell 2012), or Tactical Sound Garden (Shepard 2007). In these examples, users' interactions were stored at their location, allowing collaborations in a certain place, but separate in time.

To make music with users distributed in both space and time, some apps have allowed asynchronous layering of musical performances. Glee Karaoke (Hamilton, Smith, and Wang 2011) allows users to upload their vocal renditions of popular songs, and add layers to other performers' contributions. More conventional digital audio workstations offer social or collaboration features on mobile devices (Meikle 2016), such as uploading whole tracks or short audio clips that other users can incorporate into their compositions. Our app, MicroJam, is distinguished from these other examples by its focus on constrained and ephemeral music-making, as well as online collaboration.

MicroJam

MicroJam is an app for creating, sharing, and collaborating with tiny touchscreen musical performance

Figure 2. MicroJam allows users to browse a timeline of performances (a), create new performances (b), and reply, or play in duet, with previously recorded performances (c).

(a) (b) (c)

(Martin and Torresen 2017). It was designed to interrogate the possibilities for constrained social mobile performance. The app's design has been kept deliberately simple. The main screen (see Figure 2) recalls social-media apps for sharing images. Musical performances in MicroJam are limited to short interactions, encouraging frequent and ephemeral creative contribution. MicroJam is an iOS app written in Swift and uses Apple's CloudKit service for backend cloud storage. The source code is freely available for use and modification by other researchers and performers (https://zenodo.org/record/1412274). In this section we will discuss the design of the app, the format of the tiny musical performances, and the synthesized instruments that are available.

Design

MicroJam allows users to perform three primary activities (shown in Figure 2): to browse and listen to other users' performances, to create and share new performances using the touch screen, and to record layers on top of previously shared performances. The interface for creating new performances is in Figure 2b and is called "jam!" This screen features a square touch performance area that is initially blank. Tapping, swirling, or swiping anywhere in this area will create sounds and also start recording touch activity. All touch interaction in this area is visualized with a simple paint-style drawing that follows the user's touches. Touch interactions are simultaneously sent to a synthesized instrument to be sonified. After five seconds of touch interaction, the recording is automatically stopped (although the performer can continue to interact with the touch area). The recording can be subsequently replayed by tapping a Play icon, or it can be looped by tapping an icon in the shape of a circular arrow. Users of MicroJam can choose the instrument used to sonify their interactions in the jam interface from a button in the top right. These synthesized

instruments map a stream of touchscreen events—the location of a touch, and whether it is the beginning ("touch start") or continuation ("touch moved") of a previous gesture—to sound. The timbres and synthesis mappings are different for each instrument.

Previously recorded performances, and those recorded by other users and downloaded from the server, are listed in the World screen, as shown in Figure 2a. Each performance is represented by a visual trace of the touch drawing captured during recording, along with an indication of the contributor's online handle. Any one of these performances can be played back in place in the world screen. When playing back, both the sound and visualized touch interactions are replayed in the touch area.

When viewing a previously saved performance in the world screen, the user can tap the reply icon (a curved arrow), to open a new layer on top of the recording. As shown in Figure 2c, both the previous and current touch visualizations are shown, and each layer is sonified separately. Multiple replies in MicroJam are possible, which can result in several layers of performances being played back at once, allowing for complex compositions.

Tiny Touchscreen Performances

MicroJam is intended to provide a musical experience in which constraints are applied to users' interaction in order to increase their creativity and to lower the entry barriers for musical performance. We argued earlier that constraints in social creativity systems could actually enhance users' creative power. In the context of a musical app, these constraints could lead to more frequent interactions and possibly higher creativity owing to the lower stakes and effort. Musical interactions in MicroJam are similarly constrained to be *tiny touchscreen performances*, as they are limited in the area and duration of interaction. We define a tiny performance as follows:

1. All touches take place in a square subset of the touchscreen.

Table 1. Excerpt from a MicroJam Tiny Performance Data Record

Time	x	y	z	Moving
0.007805	0.276382	0.416080	38.640625	0
0.065901	0.286432	0.428141	38.640625	1
0.074539	0.286432	0.433166	38.640625	1
0.090817	0.290452	0.450251	38.640625	1
0.107149	0.298492	0.468342	38.640625	1
0.124072	0.309548	0.486432	38.640625	1

2. Duration of the performance is limited to five seconds.
3. Only one single touch is recorded at a time.

Such performances require little effort on the part of users. Some users may find it difficult to conceive and perform several minutes of music on a touchscreen device; five seconds is long enough to express a short idea, but short enough to leave users wanting to create another recording. It has been argued that five seconds is enough time to express a sonic object and other salient musical phenomena (Godøy, Jensenius, and Nymoen 2010). Although the limitation to a single touch may seem unnecessary on today's multitouch devices, this stipulation limits tiny performances to monophony. To create more-complex texture or harmony, performers must collaborate, or they must record multiple layers themselves.

For playback, storage, and transmission to other users, tiny touchscreen performances are recorded as simple comma-separated value files of recorded touch gestures. This data format records the user's performance movements in a compact manner (typical size is around 5KB), rather than the actual sound or abstract musical values such as notes and rests. The performance can later be recreated by sending these same touch event signals to MicroJam's synthesized instruments. The data format records each touch interaction's time (as an offset in seconds from the start of the performance), whether the touch was moving or not, x and y locations, as well as touch pressure (z), An example is shown in Table 1. The visual trace of performances is also stored as a PNG image for later use in the

Table 2. Descriptions and Mapping Details of MicroJam Instruments

Instrument	Description	x	y	dx	dy
Chirp	Basic sine oscillator sound.	Pitch	Mix lower octave	Pitch bend	
Drums	Drum sounds (bass drum, snare drum, hi-hat, and crash cymbal) triggered from each quadrant. Swipes trigger rolls.	Instrument	Instrument	Pitch bend	Reverb and delay (wet/dry mix)
FM Lead	Simple two-operator FM synthesizer that plays in a bass register.	Pitch	Volume	Pitch bend	Mix lower octave, reverb and delay (wet/dry mix)
Keys	Phase-modulation Rhodes-like keyboard sound.	Pitch	Volume	Pitch bend	Modulation depth
Pad	Sawtooth-wave pad synthesizer.	Pitch	Volume	Pitch bend	Tremolo depth, reverb and delay (wet/dry mix)
Quack	Wave packet synthesizer with analog-like sound.	Pitch	Timbre	Pitch bend	
Strings	Karplus-Strong plucked guitar synthesizer.	Pitch	Volume	Pitch bend	Reverb and delay (wet/dry mix)
Lead Guitar	Karplus-Strong synthesizer with distortion.	Pitch	Volume	Pitch bend	Reverb and delay (wet/dry mix)
Wub	Tremolo "wub wub" bass sound.	Pitch, timbre	Tremolo rate, timbre	Pitch bend	

app, although this can also be reconstructed from the touch data.

Storing gestural control data, rather than audio, or high-level musical data such as MIDI, allows performances to benefit from updated synthesis routines in the app, and future enhancements to the visual interface. As noted in previous work on preserving touchscreen improvization (Martin and Gardner 2016), this representation encodes information that might not be available in audio or MIDI recordings. In this research, we take advantage of this information to study tiny performances from the perspective of touch gesture.

Instruments in MicroJam

MicroJam includes nine different instruments (see Table 2) that map touches in the performance area to different synthesized sounds. This selection of instruments provides basic coverage of typical musical roles such as percussion (drums), bass (fmlead, wub), accompaniment (pad, keys), and lead (chirp, strings). Although far from an exhaustive collection, these instruments allow exploration of different musical ideas, both through their different timbres as well as the different touch to sound mappings used in each one. Descriptions and mapping details for each instrument is given in Table 2.

The instruments are implemented in Pure Data and make use of the rjlib library (Barknecht 2011) for some synthesis routine implementations. They are loaded in the app via libpd (Brinkmann et al. 2011). The instruments are controlled by continuous inputs of x and y values from the single touch point in performance area (z, pressure, is so far not used as it is not available in all devices). Each instrument is responsible for its specific mappings of these values to higher-level musical properties such as pitch,

timbre, and control of audio effects. For example, the FM Lead sound, a simple FM synthesizer, maps initial x values from a tap or swipe to a bass-register pitch on a continuous range (MIDI notes 36–60) without quantization to a scale or chromatic tones. The y value is mapped to volume. Tapping the screen triggers a short note, while swiping creates a sustained sound until the touch is released. For continuous swipes, the distance of the present touch point to the initial one (dx and dy) is calculated; dx is used to control pitch bend, and dy is mapped to reverb and delay effects as well as mixing in a copy of the sound in a lower octave. A similar mapping scheme is used for the other pitched sounds, see Table 2 for further details. For the Drum sound, different synthesized drum set sounds (bass drum, snare drum, hi-hat, and crash cymbal) are triggered from each quadrant of the screen, and swipes trigger a roll. In general, the instruments in MicroJam are designed to support fun audio sketches rather than song production. Constraints such as the lack of quantization on our pitched instruments are intended to limit precise note playing and to emphasize ephemeral improvisations.

Studying Tiny Performances

In this section we describe an investigation of how users interact with MicroJam through the lens of the tiny performances that have been collected in the app. Since early prototypes of MicroJam were developed, the app has been distributed and demonstrated among researchers, students, and the music technology community, and these testers and early users uploaded performances to the app's cloud database. To date, over 1,600 tiny performances are available in this database, allowing us to gain insight into the musical potential of MicroJam and the concept of tiny touchscreen performances.

Our analysis of the tiny performances is made at three levels of abstraction: individual touch events, gestures or groups of touch events, and whole performances. At the touch event level, we consider individual touch events that may be part of a swipe across the screen, or single taps. At the gesture level, we consider groups of these touch events that constitute a single "note" or interaction: either individual taps, or the collection of events that form a swipe. At the performance level, we consider descriptors of each complete five-second performance and how these vary by the instrument that was used, we also analyze the visual trace of the resulting performance. These levels reveal different aspects of the users' performance behavior; while the whole performance level demonstrates their broad artistic intentions, the gestural and touch levels expose small-scale interactions.

Participants and Data Sources

Tiny performances for this investigation came from two databases: a development database that is only accessible from instances of MicroJam installed on the authors' test devices, and a public database accessible from beta and published versions of MicroJam. Performances in the development database were made at testing and demonstration sessions taking place in our lab, at conferences, and at other events. Most of the participants in these sessions were university students and researchers in music technology or computer science. These participants had a range of music experience from untrained to professional. The public database was accessed from beta versions of the app as well as the published App Store version. Beta versions were distributed to interested members of the computer music and technology community who requested invitations through social media and at conferences. Since the public release of the app, performances in this database have been contributed anonymously by members of the public.

Unlike most music apps, all performances saved in MicroJam's public database are openly accessible through the app and through Apple's CloudKit API. Personal data (e.g., real names, e-mail addresses, IP addresses) are not included in this data. Although the use of public social-media data for research is commonplace (e.g., studies of text extracted from Twitter), there remain ethical concerns about whether users are comfortable with such use (Fiesler and Proferes 2018). In our case, MicroJam's privacy policy highlights our research intentions,

Figure 3. The visual trace of a selection of MicroJam performances with each swipe shaded differently. Some different performance styles are visible, for example, taps (dots), single long swipes (long lines or curves), repeated strokes (such as parallel lines), and drawing an image (such as a face). Some of the interactions appear to be predominantly visual, while others are exploratory scribbles with little visual coherence.
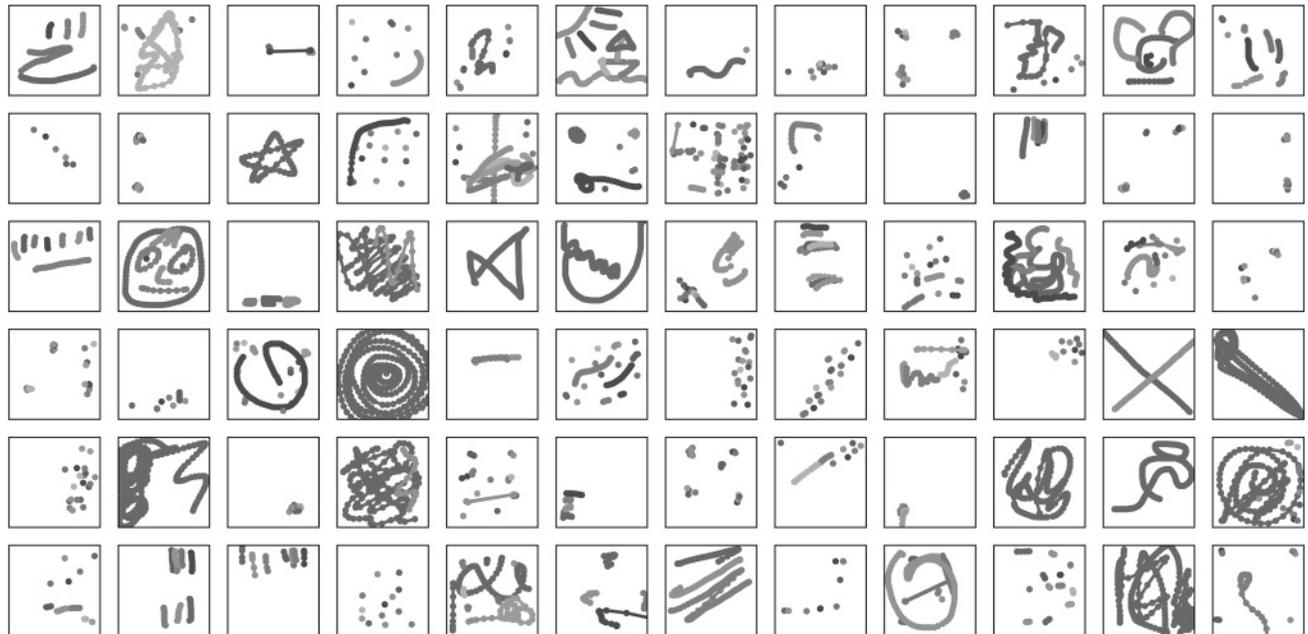
and the number of users from the published version of MicroJam is, so far, small compared with those who installed the app through private beta testing.

Within our data set, 98 unique CloudKit accounts are represented, including two accounts created for our demonstration devices. Whereas many users (including the developers) performed on our demonstration devices, we assume the other accounts are associated with individual users. The data set contains a total of 1,633 performances, of which 761 are from our demonstration accounts and the remaining 872 performances are from unidentified users. The median number of performances per user was three, although 25 percent of users created more than six performances with the app.

Performances from the first author's personal account, not used for demonstrations, were excluded from this data set. Performances on the demonstration devices, including some by the first author, were retained as these included performances from many users and further identification of individual performers was not possible.

Performance Level Investigation

The visualizations of a subset of tiny performances are reproduced in Figure 3. This figure shows the variety of styles of touch interaction generated by performers. Many of the interactions are abstract, resembling scribbles that show the user experimenting with the synthesis mapping of the jamming interface. Some performances contain patterns in which performers have repeated rhythmic motions in different parts of the touch area. A number of the performances are recognizable images: figures, faces, and words. We can characterize the visual appearance of performances under the following broad styles: taps, swipes, long swipe, mixture, image, and text. Tap and swipe performances focus on these fundamental touch gestures, whereas long swipe performances seem to consist of only one swipe, and mixture performances include mixtures of these styles. Image and text performances appear to focus on the visual meaning of the finished trace, with less emphasis apparently given to the resulting sound.

Table 3. Occurrences for Different Lengths of Reply Chain

| Chain length	2	3	4	5	6	7	8
Count	299	83	24	9	5	2	2

Of the 1,633 performances in the data set, 424 (26%) are replies. As replies to replies are possible in MicroJam, it is interesting to see how long potential chains of replies are in terms of number of performance layers (see Table 3). Of the 424, there are 299 having only two layers. Although there are 83 performances with three layers, there are scarcely any with seven or eight layers (only two for each depth). This data suggests that, although users have made some use of the reply function, they have only rarely explored its potential for creating complex, layered performances. Further development work could help encourage users to create longer performances; for instance, highlighting complex performances in the browsing screen.

Although performances with all instruments are represented in the data set, the most popular instrument is chirp (the default choice) with 322 performances. Figure 4 shows that the next five most popular instruments (Drums, Keys, Quack, Strings, and Wub) all have around 200 performances, but that FM Lead, Pad, and Lead Guitar each have fewer than 100 performances. These instruments were added to the app in later revisions, and may have been less obvious in the list of instruments.

Analysis by Touch Event

The data set includes 238,743 touch events. As set out in the tiny performance definition above, a touch event can either be a *touch start* event (when the user's finger hits the touchscreen), or a *touch moved* event (when the user's finger has moved without leaving the touchscreen). These are distinguished by a binary value, "moving," in the data set. These events are triggered by iOS touchscreen callback functions touchesBegan() and touchesMoved(), respectively. Only 13,077 touch events in the data set are nonmoving, compared with

Figure 4. Number of performances with each instrument in our data set. The default instrument, Chirp, is most popular, followed by Drums.
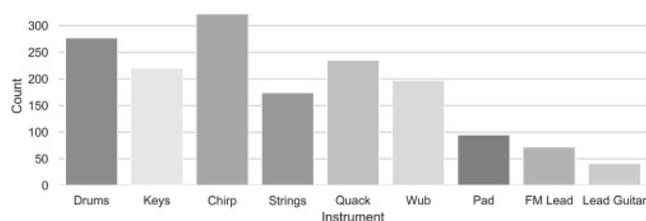
225,666 moving touches. This is because whenever a user swipes during a performance a large number of touch moves are generated, whereas only one touch start occurs.

Figure 5a shows the distribution of the number of touch events recorded per performance grouped by instrument used. This shows that for all instruments except drums, performances had a median of between 90 and 170 touch events. The median number of touch events for drums was 48, much lower than the other instruments. This is explained by Figure 5b which shows the proportion of touch events that were moving in each performance. For drums, many more touch events were taps, rather than swipes, which resulted in fewer touch events for a given performance.

Two interesting statistics are the time differences between consecutive touch events (dt), and the onscreen distance between them measured as a proportion of the performance area. Distributions of these statistics are shown in Figure 6. As expected, values of dt and distance for moving touches tend to be small (touch events in iOS are generally produced at the screen's refresh rate—60 Hz for most devices at present), although there are slower outliers which could have been caused by errors in the user interface code or drops in performance. The median dt between moving touches is only 0.017 sec (approximately 60 Hz), compared with 0.219 sec for nonmoving touches. The interquartile range of dt for nonmoving touches is 0.094–0.386 sec. This gives an indication that performers tend not to leave much time in between interactions, and that the resulting tiny performances would not have much temporal "space." Similarly, we can observe that nonmoving touches tend to have moved within a relatively small proportion of the performance area (median = 0.193, interquartile range = [0.048, 0.436]). This can be observed in some of the example performances

Figure 5. Distributions of
touch event properties
grouped by instrument
shown as letter-value plots
(Hofmann, Wickham, and
Kafadar 2017):
Distribution of the total

count of touch events in
each performance for each
instrument (a) and
distribution of moving
versus nonmoving touch
events for each instrument
(b). Each instrument had a

similar number of touch
events per performance,
with somewhat fewer for
Drums. All instruments
except Drums are
dominated by touch move
events.

Figure 6. Distributions of
touch event dt values (a)
and screen distance (b)
separated by whether the
touch was moving or not.

(a)

(b)

Figure 5

(a)

(b)
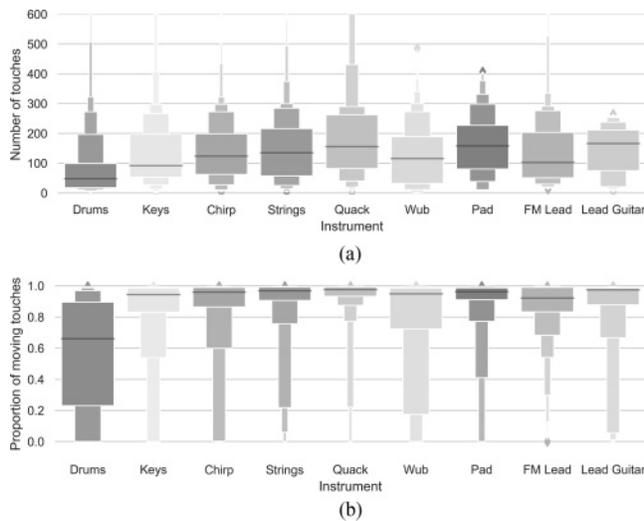
Figure 6

in Figure 3, in which a user tapped repeatedly in the same part of the performance area.

Figure 7 shows the distribution of touch events across the touchscreen performance area (obtained with a bivariate kernel density estimate). This allows us to investigate where users have most commonly interacted with the performance area to play sounds. We can observe that touches cover the whole performance area; however, more touches occur on the diagonals and in the upper right quadrant. Relatively few touches extend all the way to the edges.

These analyses suggest that users tend not to explore the potential of space in their performances, both in terms of time and the touchscreen area. Given the time limitation of five seconds, it is understandable that users would prefer to squeeze in as much activity as possible. A small number of interactions could also be effective, however, by allowing pauses that add structure to the performance (Sutton 2002). Taken together, these results could inform future synthesis mappings in the app. For instance, a mapping could produce unexpected or interesting sounds if the next touch event is far away in space or time. Given that users tend not to use the edge of the performance area, these areas could be mapped to more extreme sounds (e.g., with
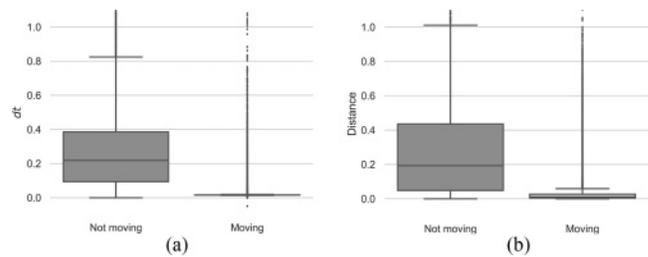
distortion or delay effects). Similar mappings have been explored in instruments such as Crackle (Reus 2011).
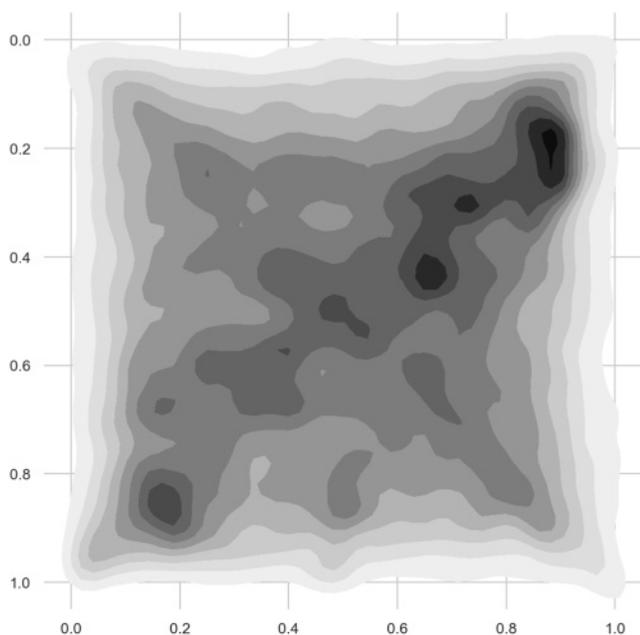
Analysis by Gesture

In this section we analyze performances from the perspective of tap and swipe gestures, the two fundamental touchscreen interactions available in MicroJam's interface. A swipe is a sequence of multiple touch events that can be defined as a touch start event followed by a nonzero number of touch moved events. The data set contains 7,237 such swipes, compared with 5,840 true taps, in which a touch start event was not followed by a touch moved event. These swipes represent one of the more important phenomena in tiny performances, as they represent the actions formed by the majority of touch points. We extracted swipes from each performance in the data set by dividing them by touch start events and discarding all divisions with only a touch start. Before analyzing swipes we removed erroneous touch events with $dt < 0.008$ sec (corresponding to the maximum 120-Hz touch scan rate in iOS), these touch events could have been caused by multiple touches being interpreted as one movement. This procedure excluded 215 swipes that were left with either one or no touch events. We excluded the 71 swipes that were longer than five seconds and were then able to perform analyses that characterize swiping behavior seen in our data set.

Table 4 shows descriptive statistics on the 7,022 valid swipes. The majority were short in time and on-screen distance (measured in proportions of

Figure 7. The distribution of touch event locations across the touchscreen performance area (bivariate kernel density estimation), with darker grays indicating a higher density of touch events. Touches broadly cover the whole performance area, but are more prominent on the diagonals and particularly in the upper-right corner.

Table 4. Statistics of Swipes in the Data Set

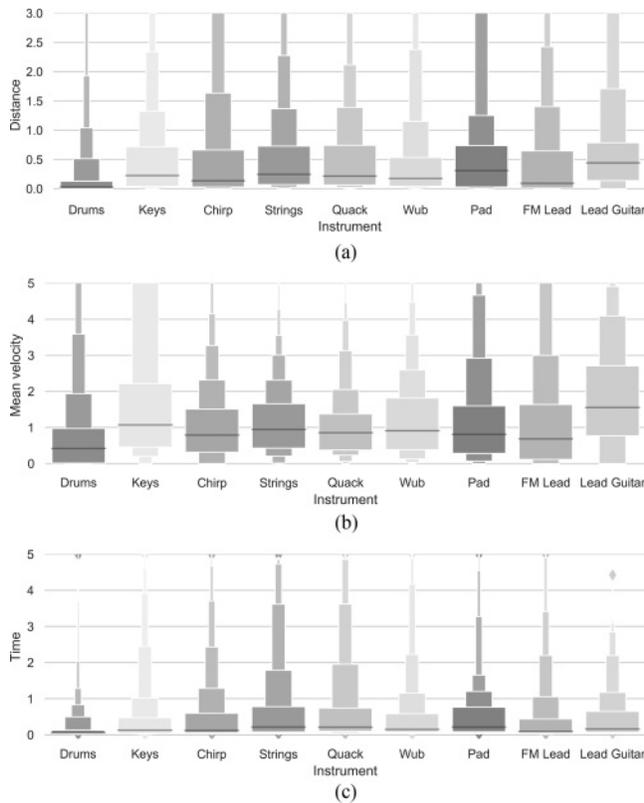|	Length	Time	Distance	Velocity Mean	Maximum
Mean	28.064	0.518	0.643	1.432	4.041
Standard deviation	52.645	0.965	1.807	2.706	8.762
Minimum	2	0.008	0.000	0.000	0.000
25%	4	0.057	0.023	0.281	0.673
50%	7	0.121	0.135	0.767	1.773
75%	26	0.506	0.590	1.574	3.743
Maximum	478	4.999	49.610	53.870	107.740

Length in number of events, time in seconds, distance in area widths, and mean and maximum velocity in area widths per second.

the performance width covered). Of the swipes, 75% were shorter than 0.506 sec in duration and the median swipe time was only 0.121 sec. The median distance was 0.135 area widths (i.e., 13.5% of the performance area) and 75% of swipes traced less than 0.59 area widths. These results seem inconsistent with the appearance of the visual traces shown in Figure 3, where the images appear to be dominated by long continuous swipes. The data set does contain long swipes, up to the maximum five-second duration and covering almost 50 area widths traversing the entire performance area (this last example is an unusual outlier). These long swipes are visually dominant, especially given that they can cover up some smaller interactions, but they are outnumbered by the shorter swipes that make up the vast majority of gestures. A small number of swipes seem to have incredibly fast movements. These are due to instances where multiple touches on the screen have been erroneously interpreted as movement.

In Figure 8, we show the distributions of time, distance, and mean velocity for each instrument in MicroJam. One-way ANOVA tests on each measurement confirm that the instrument used

has a statistically significant effect ($p < 0.001$) on these attributes. In particular, Drum performances have much shorter swipes than any of the other instruments in terms of both distance traced and time. This could be due to many attempted taps that were measured as short swipes with just a few touch points. The Keys instrument seemed to gather some of the fastest swipes.

To gain an intuitive idea of what these swipes looked like, we have visualized selections of swipes of different durations. These are shown in Figure 9. First, the quartiles for the time dimension were calculated (shown in Table 4), then 200 swipes were randomly sampled from each quartile. Swipes in the first quartile were almost always straight lines with some spanning quite far across the screen. The two quartiles around the median length showed much variation in expression. Swipes shorter than the median rarely had more than a subtle curve, and those above the median showed curves that could have an expressive effect on the pitch and timbre of the resulting sounds. In the upper quartile of length, swipes could cover the whole performance area or trace complex patterns. These swipes could represent longer notes, parts of drawings, or entire five-second performances.

The 2-D traces of performances in Figures 3 and 9 do not show the velocity of swipes, a quantity with much expressive potential. In Figure 10 we visualize the normalized velocity curves for selections of swipes of different lengths. The shortest swipes
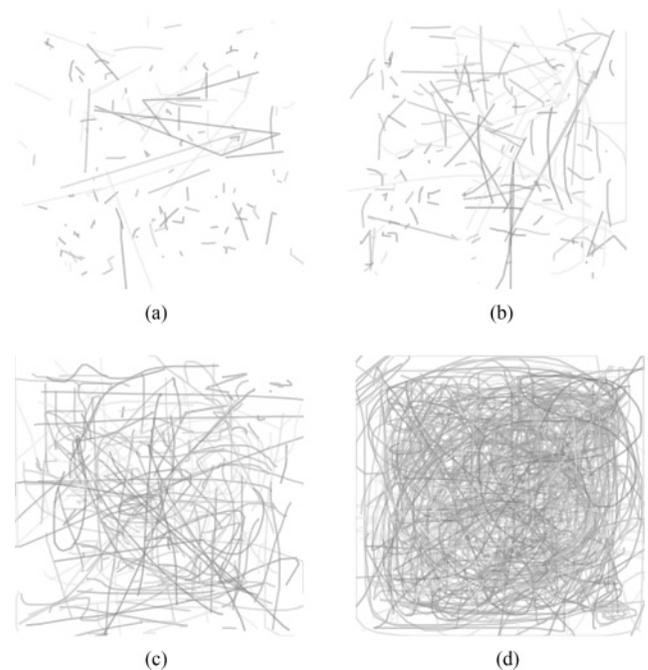
Martin and Torresen **51**

Figure 8. The distribution of swipe statistics for different instruments shown as letter-value plots (Hofmann, Wickham, and Kafadar 2017): Distance in area widths (i.e., fraction of the width of the performance area) traveled over each swipe (a), mean velocity in area widths per second over each swipe (b), and time in seconds taken for each swipe (c).

Figure 9. Random selections of 200 swipes from each quartile of swipe-length duration represented in the data set (quartiles displayed in order from (a) to (d)).

have only two or three touch points, and the velocity tends to only increase, indicating a quick flicking movement. The next quartile shows a more expressive curve, with a quick rise and slower release as the touch point stops moving before the end of the swipe. The third quartile shows the possibility for multiple peaks and valleys in the velocity, perhaps indicating changes in direction of the moving touch point. Again, only the fourth quartile shows extensive expressive behavior, such as repeating peaks in velocity that could indicate a rhythmic movement pattern over a longer swipe.

The analysis of swipes in MicroJam provides some important insights. Most importantly, the majority of swipes are short—three quarters are less than 0.506 seconds. The longest swipes have more scope for expressive behavior, however. This has implications for future instrument design that should allow more detailed expression with short interactions, e.g., by slowing down the effect a

swipe has on audio effects or timbral changes to make them more noticeable. Alternatively, long swipes could be encouraged. It could be that, if the visual impact of long swipes were reduced by fading them out over time, users might feel more confident to explore long swipes more frequently. Further refinements to instruments, such as Drums, that reward longer swipes with interesting sounds (e.g., pitch bends on toms or cymbals) could also encourage such behavior.

Discussion

The results of our study allow us to characterize how users perform within the constraints of MicroJam's tiny performance idiom. We know that they perform in different broad styles including abstract gestures as well as images and text. Although visually meaningful performances draw the eye, the data set is dominated by taps and the shortest of swipes, which are typical of performance styles that are more abstract. It is questionable whether image and text explorations lead to rewarding musical experiments,

Figure 10. The normalized velocity curves of random selections of 20 swipes from each duration quartile represented in the data set. The shortest swipes tend to be strictly increasing in velocity whereas the two longer quartiles show some temporal activity, such as pulsing velocity, that could indicate tracing back and forth across the screen.
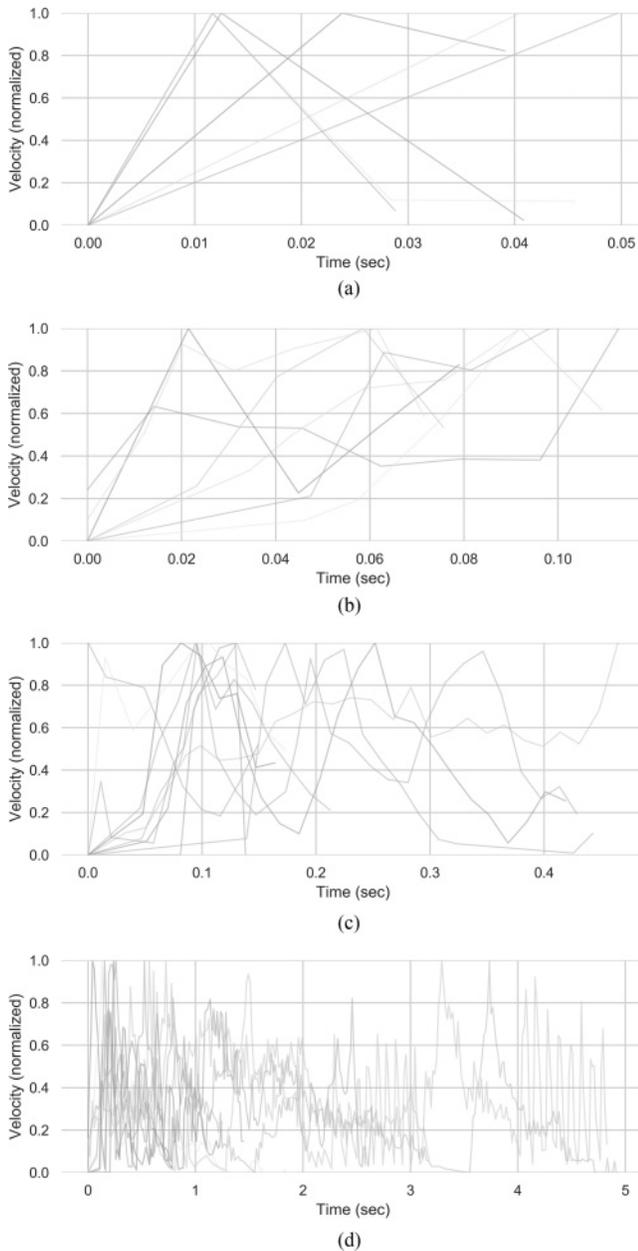
which limits the conclusions we can draw. Future revisions of MicroJam could emphasize replying and collaboration rather than just performance creation.

Four of our findings from analysis of MicroJam's touch data can be extrapolated into design recommendations for future versions of MicroJam and other touchscreen music systems:

1. At present, few performances have more than two layers. To encourage performances that are more complex, multilayered performances could be celebrated within the app. Multilayered performances could be highlighted in the world feed, and opportunities to reply presented more actively.
2. The edges of the performance area are rarely used by performers. The instrument mappings could be altered to use these spaces to create more exotic or experimental sounds that reward the user's exploration.
3. Very short swipes are the most common gesture. To emphasize the impact of these interactions, they could be emphasized more in the sound design for MicroJam.
4. Long swipes are rare, yet they dominate the visual trace of performances. The visual impact of long swipes could be reduced, for instance, by fading out the long tails of interactions. This might allow multiple instances of them to be used without overwhelming the touch area.

So far, MicroJam has mainly been used in test and demonstration environments, and few users have shared large numbers of performances. As a result, the performances analyzed in this data set are generally by inexperienced users. We would expect, however, that as for other instruments, MicroJam users would improve with practice, and develop new styles. Future work could seek to identify changes in tiny performance style over time.

One aspect of analysis that has not been mentioned is modelling and generation of tiny performances with machine learning. Previous research has already discussed a mixture density, recurrent neural network model for generating and responding to tiny performances (Martin and Torresen 2018).

and it may be more appropriate to focus on improving the expressive potential of other performance styles. As for the social aspects of MicroJam, although the reply function has certainly been used in the data set, few multilayered performances are present,

Martin and Torresen **53**

This system is available in the app to provide an automatic reply to a performance on demand, and it generates the same control data format as the tiny performances. In future work we could explore the potential for developing tailored models of individual users' styles which could even provide control over the kind of gestures (for instance, short swipes or taps) that are provided in an automatic response. These models could also be used to demonstrate effective use of the gestures highlighted in the above recommendations as part of a performance training feature. Future studies could also compare longer performances with MicroJam to the present five-second format to confirm whether the constrained performance length affects performers' interactions.

Conclusions

In this article, we have presented the design for MicroJam, a social mobile music app for creating touchscreen performances, and defined the tiny performance format. We have investigated this app through a data-driven analysis of more than 1,600 performances created by more than 97 individual users. This investigation has revealed how users perform in the tiny touchscreen idiom and revealed contradictions and complexities in these performances. We now know that short swipes are the most common gesture, yet long swipes dominate the visual appearance of performances. We found that the edges of the performance area are not well explored by performers and that users do not explore space and time as much as they could. Still, the visual representations in performances are varied and expressive. These findings enhance our understanding of how users interact with free-form touchscreen music systems and could apply to a variety of app designs.

MicroJam is an example of a social media app centered on musical creation rather than written and visual media. We have argued that such apps could take advantage of the ubiquity of mobile devices by allowing users to collaborate asynchronously and in different locations, and shown that these modes of interaction are relatively unexplored compared to more conventional ensemble performances. The

focus on time-limited tiny performances represents a new approach to asynchronous musical collaboration. Taking inspiration from the constrained contributions that typify social media apps, MicroJam limits users to five-second touchscreen performances, but affords them extensive opportunities to browse, to play back, and to collaborate through responses. MicroJam's tiny performance format includes a complete listing of the touch interactions and so allows performances to be easily distributed, visualized, and studied.

Our novel data-driven investigation examined 1,633 performances consisting of 238,743 touch events. Our quantitative method sets this work apart from previous research that has focussed on users' qualitative experience. The performances were analyzed at the levels of individual touch events, grouped touch gestures, and whole performances. The investigation revealed the variety of styles used in performances but that fewer performances than desired were replies. Examining touch points showed that the edges of the performance area was not used as much as the center and main diagonals, and that moving touches, rather than taps, dominated the data set. Grouping touches into swipes showed that while long swipes are more visually apparent, the vast majority of swipe gestures are actually short. This data-driven method is a contribution of this work, and could be applied to other studies of musical performance systems on touchscreens and other interfaces.

The analysis in this article has suggested that even a simple and constrained touchscreen interface can lead to a variety of styles and unexpected musical interactions. Although constraining the length of performances may have both increased the number recorded and made it easier to collaborate using musical replies, it could also have curbed users' gestural exploration with the touch screen. We have discussed how our findings might inform design revisions that could better align MicroJam's capacity for musical expression with user behavior. These revisions may encourage more-expressive performances from MicroJam users, and more collaboration among them, without increasing the effort required to generate performances with a tiny touchscreen. For music-making, as opposed

to appreciation, to be widely adopted as part of everyday social media interactions, this balance between constraint and expression will need to be further examined and addressed. We posit that a data-driven approach to mobile music performance, examining musical data generated by users, can be used to further examine this balance in MicroJam and other systems for mobile musical creativity.

Acknowledgments

References

Allison, J., and C. Dell. 2012. "AuRal: A Mobile Interactive System for Geo-Locative Audio Synthesis." In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Available online at www.nime.org/proceedings/2012/nime2012_301.pdf. Accessed June 2020.

Barknecht, F. 2011. "rj: Abstractions for Getting Things Done." In *Proceedings of the Pure Data Convention*, pages unnumbered.

Barraclough, T. J., D. A. Carnegie, and A. Kapur. 2015. "Musical Instrument Design Process for Mobile Technology." In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 289–292.

Behrendt, F. 2012. "The Sound of Locative Media." *Convergence* 18(3):283–295.

Brinkmann, P., et al. 2011. "Embedding Pure Data with libpd." In *Proceedings of the Pure Data Convention*. Available online at www.uni-weimar.de/medien/wiki/PDCON:Conference/Embedding_Pure_Data_with_libpd:_Design_and_Workflow. Accessed June 2020.

Bryan-Kinns, N. 2004. "Daisyphone: The Design and Impact of a Novel Environment for Remote Group Music Improvisation." In *Proceedings of the Conference on Designing Interactive Systems*, pp. 135–144.

Carôt, A., P. Rebelo, and A. Renaud. 2007. "Networked Music Performance: State of the Art." In *Proceedings of the Audio Engineering Society Conference*. Available online at www.aes.org/e-lib/browse.cfm?elib=13914 (subscription required). Accessed June 2020.

Correa, A. G. D., et al. 2009. "Computer Assisted Music Therapy: A Case Study of an Augmented Reality Musical System for Children with Cerebral Palsy Rehabilitation." In *IEEE International Conference on Advanced Learning Technologies*, pp. 218–220.

d'Alessandro, N., et al. 2012. "A Digital Mobile Choir: Joining Two Interfaces towards Composing and Performing Collaborative Mobile Music." In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Available online at www.nime.org/proceedings/2012/nime2012_310.pdf. Accessed June 2020.

Dewan, S., and J. Ramaprasad. 2014. "Social Media, Traditional Media, and Music Sales." *Management Information Systems Quarterly* 38(1):101–121.

Essl, G., and S. W. Lee. 2018. "Mobile Devices as Musical Instruments: State of the Art and Future Prospects." In M. Aramaki et al., eds. *Music Technology with Swing*. Berlin: Springer, pp. 525–539.

Essl, G., and M. Rohs. 2009. "Interactivity for Mobile Music-Making." *Organised Sound* 14(2):197–207.

Favilla, S., and S. Pedell. 2013. "Touch Screen Ensemble Music: Collaborative Interaction for Older People with Dementia." In *Proceedings of the Australian Computer-Human Interaction Conference*, pp. 481–484.

Fiesler, C., and N. Proferes. 2018. " 'Participant' Perceptions of Twitter Research Ethics." *Social Media and Society* 4(1). Available online at journals.sagepub.com/doi/pdf/10.1177/2056305118763366. Accessed June 2020.

Gaye, L., et al. 2006. "Mobile Music Technology: Report on an Emerging Community." In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 22–25.

Gligorić, K., A. Anderson, and R. West. 2018. "How Constraints Affect Content: The Case of Twitter's Switch from 140 to 280 Characters." In *Proceedings of the International AAAI Conference on Web and Social Media*, pp. 596–599.

Godøy, R. I., A. R. Jensenius, and K. Nymoen. 2010. "Chunking in Music by Coarticulation." *Acta Acustica united with Acustica* 96(4):690–700.

Greenberg, S., and M. Roseman. 1998. "Using a Room Metaphor to Ease Transitions in Groupware." Technical Report 98/611/02. Canada, Department of Computer Science, University of Calgary.

Gurevich, M., A. Marquez-Borbon, and P. Stapleton. 2012. "Playing with Constraints: Stylistic Variation with a Simple Electronic Instrument." *Computer Music Journal* 36(1):23–41.

Hamilton, R., J. Smith, and G. Wang. 2011. "Social Composition: Musical Data Systems for Expressive Mobile Music." *Leonardo Music Journal* 21:57–64.

Hofmann, H., H. Wickham, and K. Kafadar. 2017. "Letter-Value Plots: Boxplots for Large Data." *Journal of Computational and Graphical Statistics* 26(3):469–477.

John, D. 2013. "Updating the Classifications of Mobile Music Projects." In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 301–306.

Jordà, S. 1999. "Faust Music on Line (FMOL): An Approach to Real-Time Collective Composition on the Internet." *Leonardo Music Journal* 9:5–12.

Kelkar, T., and A. R. Jensenius. 2018. "Analyzing Free-Hand Sound-Tracings of Melodic Phrases." *Applied Sciences* 8(1):35–55.

Kim-Boyle, D. 2009. "Network Musics: Play, Engagement and the Democratization of Performance." *Contemporary Music Review* 28(4–5):363–375.

Martin, C., and H. Gardner. 2016. "A Percussion-Focussed Approach to Preserving Touch-Screen Improvisation." In D. England, T. Schiphorst, and N. Bryan-Kinns, eds. *Curating the Digital: Spaces for Art and Interaction*. Berlin: Springer, pp. 51–72.

Martin, C., H. Gardner, and B. Swift. 2015. "Tracking Ensemble Performance on Touch-Screens with Gesture Classification and Transition Matrices." In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 359–364.

Martin, C., et al. 2016. "Intelligent Agents and Networked Buttons Improve Free-Improvised Ensemble Music-Making on Touch-Screens." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2295–2306.

Martin, C. P., and J. Torresen. 2017. "Exploring Social Mobile Music with Tiny Touch-Screen Performances." In *Proceedings of the Sound and Music Computing Conference*, pp. 175–180.

Martin, C. P., and J. Torresen. 2018. "RoboJam: A Musical Mixture Density Network for Collaborative Touch-screen Interaction." In A. Liapis, J. J. Romero Cardalda, and A. Ekárt, eds. *Computational Intelligence in Music, Sound, Art, and Design*. Berlin: Springer, pp. 161–176.

Meikle, G. 2016. "Examining the Effects of Experimental/Academic Electroacoustic and Popular Electronic Musics on the Evolution and Development of Human–Computer Interaction in Music." *Contemporary Music Review* 35(2):224–241.

Oh, J., et al. 2010. "Evolving the Mobile Phone Orchestra." In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 82–87.

Redi, M., et al. 2014. "Six Seconds of Sound and Vision: Creativity in Micro-Videos." In *The IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4272–4279.

Reus, J. 2011. "Crackle: A Mobile Multitouch Topology for Exploratory Sound Interaction." In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 377–380.

Scheffel, S., and B. Matney. 2014. "Percussion Use and Training: A Survey of Music Therapy Clinicians." *Journal of Music Therapy* 51(1):39–63.

Schiemer, G., and M. Havryliv. 2007. "Pocket Gamelan: Swinging Phones and ad-hoc Standards." In *Proceedings of the International Mobile Music Workshop*. Available online at ro.uow.edu.au/creartspapers/24. Accessed July 2020.

Schlei, K. 2012. "TC-11: A Programmable Multi-Touch Synthesizer for the iPad." In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Available online at www.nime.org/proceedings/2012/nime2012_230.pdf. Accessed June 2020.

Shepard, M. 2007. "Tactical Sound Garden Toolkit." In *ACM SIGGRAPH 2007 Art Gallery*, SIGGRAPH '07, p. 219.

Smith, J. C. 2013. "Correlation Analyses of Encoded Music Performance." PhD dissertation, Department of Music, Stanford University.

Stokes, P. D. 2008. "Creativity from Constraints: What Can We Learn from Motherwell? From Modrian? From Klee?" *Journal of Creative Behavior* 42(4):223–236.

Sutton, J. P. 2002. " 'The Pause That Follows'. . . Silence, Improvised Music, and Music Therapy." *Nordic Journal of Music Therapy* 11(1):27–38.

Swift, B. 2013. "Chasing a Feeling: Experience in Computer Supported Jamming." In S. Holland et al., eds. *Music and Human–Computer Interaction*. Berlin: Springer, pp. 85–99.

Tanaka, A. 2010. "Mapping Out Instruments, Affordances, and Mobiles." In *Proceedings of the International*

Conference on New Interfaces for Musical Expression, pp. 88–93.

Trump, S., and J. Bullock. 2014. "Orphion: A Gestural Multi-Touch Instrument for the iPad." In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 159–162.

Wang, G. 2014. "Ocarina: Designing the iPhone's Magic Flute." *Computer Music Journal* 38(2):8–21.

Wang, G., J. Oh, and T. Lieber. 2011. "Designing for the iPad: Magic Fiddle." In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 197–202.

Wang, G., et al. 2015. "World Stage: Crowdsourcing Paradigm for Expressive Social Mobile Music." *Journal of New Music Research* 44(2):112–128.