Annotating Distributed Scores for Mutual Engagement in Daisyphone and Beyond

Nick Bryan-Kinns

ABSTRACT

Written and drawn annotations of musical scores form a core part of the music composition process for both individuals and groups. This article reflects on the annotations made in new forms of distributed music-making wherein the score and its annotations are shared across the web. Four kinds of annotation are identified from 8 years of studies of mutual engagement through distributed music-making systems. It is suggested that new forms of web-based music-making might benefit from shared and persistent graphical annotation mechanisms.

Whilst music has an important social role [1] and aural tradition [2], graphical sketches and written notes have formed a key part of many composers' [3] and musical groups' [4] practice for generations. These materials support idea generation, compositional structuring, communication, discussion and reflection for both groups and individuals. With the advent of computer-mediated music-making, especially when groups of musicians are no longer co-present, graphical and textual annotations may take on new roles in managing the creative process. This article presents observations on the emergent use of sketches and textual annotations in distributed music-making where the score, music and any additional information is shared across the Internet. The focus here is not on attempting to replicate naturalistic group music-making interaction across the Internet using high-quality video links [5] or other high-bandwidth communication channels but rather on exploring what can be learnt from studies of communication in naturalistic settings and exploring the design of low-bandwidth collaborative music systems, which I call *distributed music-making* systems. Here I report on the development, study and refinement since 2003 of a specific set of distributed music-making tools, starting with the Daisyphone system.

Nick Bryan-Kinns (researcher, teacher), Queen Mary University of London, Centre for Digital Music and Interaction, Media and Communications Research Group, School of Electronic Engineering and Computer Science, Mile End, London E1 4NS, U.K. E-mail: <nickbk@eecs.qmul.ac.uk>.

See <mitpressjournals.org/toc/lmj/-/21> for supplemental files (such as audio and video) related to this issue of LMJ and accompanying CD. See also <Daisyphone.net> for materials related to this article.

DISTRIBUTED MUSIC-MAKING

Föllmer [6] and Blaine and Fels [7] conducted early surveys of collaborative music-making systems and their origins, highlighting key characteristics ranging from the media used to the systems' learning curves and physicality. Föllmer's survey included tracing the origins of networked music in live performance, sound toys, academic research, media artists and technologically driven developments. Relying on extensive studies of "Net Music" systems, Föllmer also characterized systems in terms of "interplay with network characteristics," "interactivity/openness" and "complexity/flexibility" [8]. Distributed music systems do not attempt to replicate face-to-face interaction and so, in Föllmer's terms, may embrace or even actively exploit network characteristics such as latency. Moreover, they tend to be open to a diverse

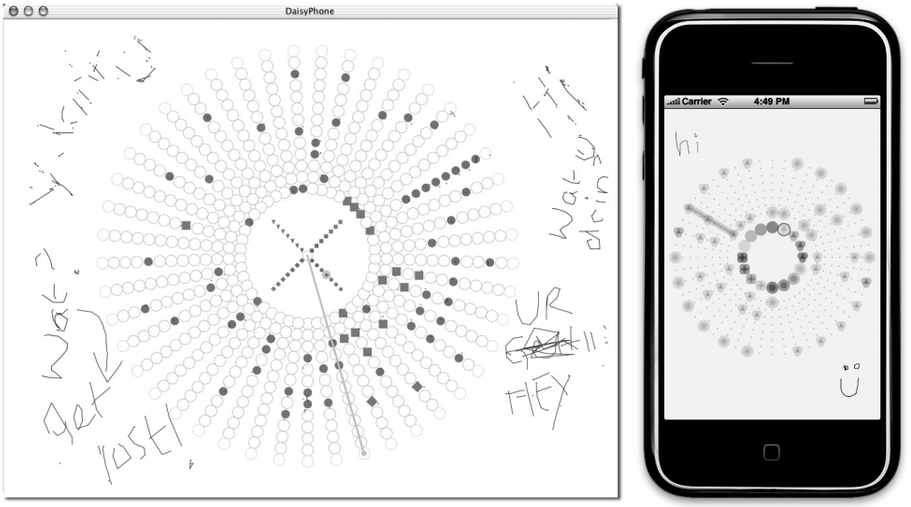


Fig. 1. Daisyphone in use on a desktop PC (left) and an Apple iPhone (right). (© N. Bryan-Kinns)

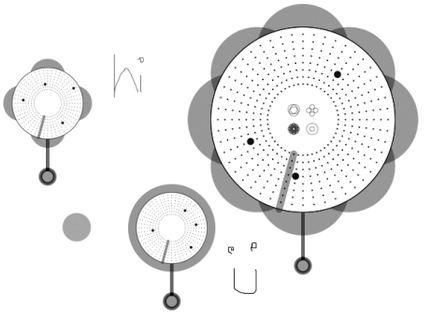


Fig. 2. Daisyfield user interface with three loops of music and graphical annotation. (© N. Bryan-Kinns)

Fig. 3. Annotations about presence—written names and questions about who is using Daisy*. (© N. Bryan-Kinns)

set of participants and range from quite simple to complex systems for musical expression. Since the work accomplished by Föllmer, and by Blaine and Fels, there has been an increase in interest in co-located shared musical experiences, such as the reacTable [9], and distributed music-making systems such as Ocarina [10], but we still know relatively little about how people go about using these systems. Barbosa captured this increase through a thorough survey of the field in 2006 [11] and developed the Networked Music Classification Space to distinguish between local and remote physical location of participants and between synchronous and asynchronous interaction. Barbosa's work also provided extensive coverage of different forms of networked music systems, from distributed composition tools to real-time performance systems, and explored the implications that the nature of networked computers has for the possible forms of music-making available. Weinberg's survey of the field [12] instead focused on the social imperative of networked music-making and on how live performance could be supported through technologically mediated interaction or interconnected music networks. More recently, Mills [13] provided a contemporary survey of networked music-making systems, focusing on software platforms for music collaboration as opposed to individual systems, reflecting the increasing maturity of the field.

What is particularly interesting about distributed music-making is that all communication must be mediated through the system, unlike with co-located situations, in which participants can see, hear and possibly feel each other [14]. Therefore, making music across the Internet with others typically involves at least some shared visual representation of the musical score being composed or contributed to. The question I explore here is: What patterns of use emerge with such shared representations?

From Daisyphone to Daisyfield

In October 2003 I launched the distributed music-making tool called Daisyphone [15–17]. At launch it received between four and 18 players per day from all over the world. Logs of interaction have been collected since its launch, and there are now 160Mb of log files, or approximately 10 million individual interactions.

Daisyphone provides a shared loop of music (5 seconds; 48 beats) that can be edited by up to 10 people at once. In this distributed and shared score, each participant can create notes using four different voices and can edit any notes in the score. Shared graphical annotation is also supported around and within the musical loop, or *Daisy*. The left image of Fig. 1 illustrates Daisyphone in use on a desktop PC. The score is represented by the circle of dots, and the currently played set of notes is indicated by the grey line radiating from the center, which rotates clockwise over a period of 5 seconds. Shapes represent different kinds of sound (ambient electronic sound palette in C-major scale consisting of bass, lead, wash and percussion), and the shapes in the center allow participants to select the sound type and volume with which they create notes. Colors are uniquely assigned to users to provide a simple sense of identity. In essence, Daisyphone allows a small group of participants to co-create musical scores that are shared, edited and performed in quasi-real time—participants hear the same audio produced from the shared score but they might not be listening to exactly the same point in the score at the same time (but remember that the score is only 5 seconds long), and the score is updated in real time as participants edit it, so network delays might mean that changes do not happen at exactly the same time for each participant. The short length of the loop combined with the quasi–real-time synchronization and the persistence of the score means that this form of distributed music-making is somewhat novel—a form of on-line improvised composition. From a user-experience perspective, this is slightly different from what Barbosa [18] refers to as Remote Synchronous Network Music systems, as the client-server updates may not happen immediately or synchronously, meaning that participants hear loops played at different times and possibly slightly different scores for short periods of time as the updates propagate across the network—it is more akin to Remote *semi*-Synchronous Network Music. The score and graphical annotation sharing is supported using a client-server architecture similar to earlier work on WebDrumII [19].

Early versions of Daisyphone ran on web browsers using Java on desktop PCs and tablet PCs. In February 2009 a version of Daisyphone for the Apple iPhone was launched as illustrated in the right-hand image of Fig. 1. In the iPhone version, the interface has less space for annotations and a revised sound selector interface but uses the same networking protocols, meaning that people can interact with each other using web browsers and iOS devices. Since its launch, Daisyphone has been used in a wide range of environments, from controlled experiments to serendipitous public uses, in exhibitions [20] and at live concerts.

In 2009 I developed a richer version of Daisyphone—Daisyfield—that allows

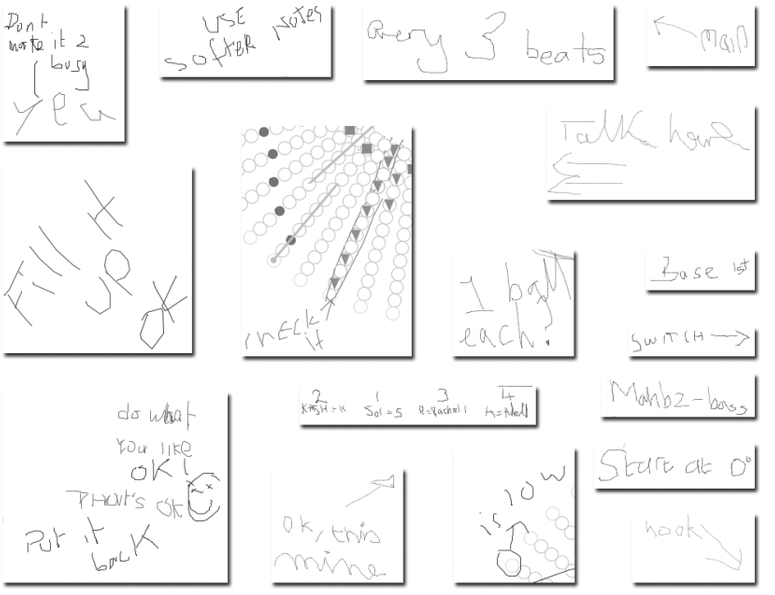


Fig. 4. Making it hSappen: annotations to manage the creative process. (© N. Bryan-Kinns)

multiple loops (i.e. Daisys) to be played concurrently and for participants to arrange Daisys in a two-dimensional space that can be graphically annotated. Again, a single shared score of 48 beats is created from the individual Daisys placed in the space, and each participant hears the same audio output. Figure 2 illustrates the Daisyfield interface with three Daisys shown, one being open for editing. The aim of this interface is to provide a richer musical and communicative user interface for exploring distributed music-making.

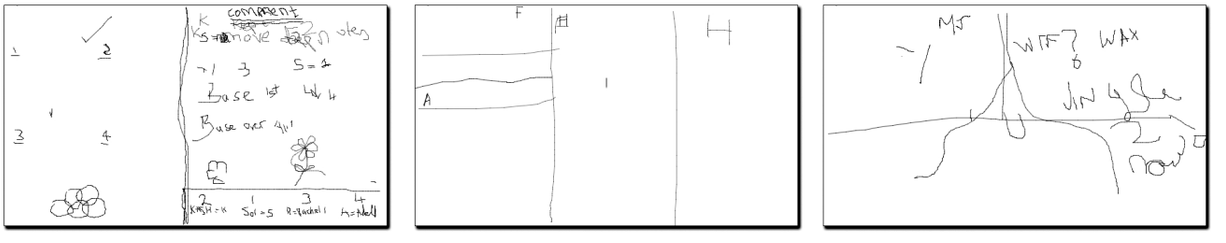
MUTUAL ENGAGEMENT

From studies of Daisyphone and Daisyfield (collectively referred to as *Daisy**), I have observed a number of similar patterns of graphical annotations when participants *mutually engage* [21] in distributed music-making, which I discuss below. Mutual engagement refers to participants engaging with each other *and* with the product they are jointly creating—focused moments of group flow [22]. My previous research has explored how different cues to mutual awareness and different features of the music-making system affect mutual engagement; for example, controlled experiments show that providing cues to participants' identity increases their mutual engagement, and post-hoc observations and categorizations of participants' activity indicated that providing persistent musical scores enables learning [23]. The controlled experiments involved participants creating short loops together over the period of an hour, with systematic changes to the user interfaces that were correlated with measures of interaction including counting numbers of musical contributions and mutual editing of each others' contributions, as well as post-task questionnaires completed by participants. The key differentiator of Daisy* to co-located shared music experiences is that the activity of music-making is completely mediated through the shared visual interface. As such, participants are simultaneously undertaking some aspects of the roles of performer, composer and audience of their collective piece of music but, curiously, they may not know the actual identity of the other participants. (I am purposely steering away from theoretical discussion of the distinctions between performer, composer and audience in such situations to focus on the forms of interaction that emerge. I leave the former topics to the reader to consider.) I refer here to users of Daisy* as *participants*—they interact somehow with the technology and each other in the process of making and appreciating music as a group, or musicking [24]. The shared graphical annotations of Daisy* then become both the instructions for the composition and the medium of social and work organization, as I discuss below.

ANNOTATIONS

Whilst it has been shown that people can create music using Daisy* without graphical annotations [25] (i.e. purely through the shared score), participants appear to be much more mutually engaged when they have an additional communication channel, as indicated by participants' responses to post-study interviews and questionnaires. The interesting question is: What do people communicate through shared annotations when they are engaged in distributed music-making? Through previous studies [26], and drawing on the work of Computer Supported Cooperative Work researchers [27] (particularly on workspace awareness—participants' ongoing understanding of who has done what and where), I identified five categories of use of shared annotations with Daisy*: Presence, Making It Happen, Quality, Social and Localisation. These categories are illustrated here using examples drawn from the Daisy* corpus of logfiles. Whilst CSCW research tends to be very workplace-oriented in its analysis of interaction, resulting in categorizations of collaborators' discussion topics in terms of design issues, project management, meeting management, etc., the CSCW research process of identifying catego-

Fig. 5. Three kinds of screen division in Daisyfield used to coordinate music-making. (© N. Bryan-Kinns)



ries of textual discussion can usefully be applied to understanding distributed music-making systems.

Presence

Daisy* does not provide any indication of who is present in the collaboration, so unsurprisingly, quite a lot of initial annotations in a composition session revolve around finding out who is in the session and making others aware of one's presence. Figure 3 illustrates some example statements of presence (e.g. "ITZ ME FAZ") and queries about who is present (e.g. "Anyone here?"). As participants have unique colors, the statements of presence serve both as indicators of who is collaborating and also to identify future contributions as belonging to a specific contributor. Finding out who is also present in Daisy* is particularly important for serendipitous web-based use where chance encounters lead to music-making.

Making It Happen

After identifying themselves to each other, participants usually get down to the task of making music together. Again, as Daisy* does not force any roles on users, it is up the participants to organize themselves somehow. Figure 4 illustrates some of the annotations participants use to coordinate their activity, from score-oriented instructions such as "Fill it up" to identifying the purpose of parts of the score, e.g. "main," and task division between participants, e.g. "1 ball each?" (referring to the loops in Daisyfield). Again, the free-form nature of the shared annotations in Daisy* allows for different groups to work in ways that suit them—dynamically reconfigurable task management emerges from the participants and is not enforced by the system. This is particularly well illustrated by the use of Daisyfield's larger shared annotation area to divide the space into areas for different activities, as illustrated in the three screenshots of Fig. 5. The left-most screenshot shows participants dividing their shared space into composition (left) and discussion areas (right); the center screenshot shows participants dividing their space into three columns (one for each participant), and the right-hand screenshot shows division of the shared space into three from the center point. This behavior illustrates the importance of providing large shared annotation spaces for participants to construct their own instructions for management of the composition of shared scores. It is interesting to note that participants very rarely use the graphical annotations as a way of sketching out musical ideas but rather use the space to coordinate the process of composition.

Quality

With the process of making music underway, numerous comments about the quality of the music being produced, e.g. "this sounds awesome" and "do u like it?," as illustrated in Fig. 6, can be observed in controlled studies, observations of public events and logs of on-line usage. These emerging discussions are critical to the success of the collaboration and provide a mechanism for participants to reflect on their music-making and to engage in quality judgments—moving the focus of the discourse to the key topic of making good music together.

Social

Aside from actually making the music and managing their on-line presences, participants routinely engaged in social, or non-task related, discussions. Figure 7 illustrates these words or images unconnected to any of the categories of annotation above—for example, drawings of faces, or "LOL" not directly connected to any parts of the shared score. Many of these annotations were used in humorous ways, e.g. drawing a funny-looking fish or smiley faces. These light-hearted contributions form the social glue of the interaction rather than contributing to the musical score per se.

Localisation

Across the four categories of textual annotations outlined above, a key use of the shared annotations in Daisy* is localization—identifying important areas, parts and words and connecting them together. Typically this is achieved by drawing arrows as illustrated in Fig. 4—one participant has written "check it" with an arrow pointing to a highlighted part of the score. Similarly, in Fig. 6 a participant has connected a sad face to a particular piece of the score to indicate his or her feelings about it. These forms of localization are critical to mutually engaging refinement of shared scores—they allow participants to directly embed their thoughts in the shared representation and highlight pertinent parts of the shared score. In other words, the ability to localize comments to musical objects supports the work of group music-making.

Reflections and Summary

The forms of shared annotations outlined in this article embody the communication that I have observed emerging around shared and distributed musical scores. It is interesting that similar patterns of communication emerge on the web and on mobile platforms and in the range of different contexts, from controlled studies to public exhibitions. Moving from the Daisyphone to the Daisyfield, I have found that the size of the resultant drawing is the only significant change in the use of annotation, with Daisyfield's larger drawing area being used to coordinate the task of distributed music-making more than Daisyphone's. However, it is worth noting that the Daisy* systems are very simple networked music systems that do not pro-

Fig. 6. Annotations about the quality of music being made. (© N. Bryan-Kinns)

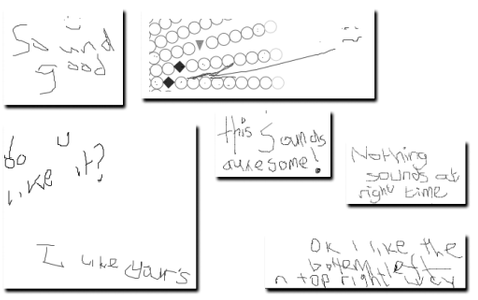


Fig. 7. Social annotations that do not concern making music. (© N. Bryan-Kinns)

vide the rich and immersive group music experience of tele-present or co-located systems, partly to enable focus on what happens when rich human-interaction channels are not available. Where other communication channels are available, the use of shared annotations may return to the function of "annotation for later review" rather than annotation for real-time interaction. It would be interesting to explore how shared annotation might be deployed usefully in richer environments.

In designing future web-based music-making systems, shared annotation mechanisms that support task management, presence awareness, social interaction and, particularly, localization, seem to be important. A conventional text chat system would struggle to support all these mechanisms in a fluid and engaging way—especially localization of annotations—as text may appear in a text window separate from the score itself. This has implications for systems that Barbosa refers to as Networked Music Systems, both synchronous and asynchronous. Tools such as Ocarina provide cues to identities of creators but do not explicitly support collaborative editing of distributed scores. Moreover, although they provide a "like" tag to allow people to comment on the quality of one another's compositions, there is little social interaction supported. Similarly, interactive tabletop interfaces such as reacTable can be linked together to create distributed music-making environments, providing a Local Networked Music experience, but they provide no annotation mechanisms, even though the physical form of interactive tabletops (resembling conventional tables and desks) provide an intuitive drawing area. It would be interesting to explore future versions of such systems that combine their playful music-making interfaces with richer shared editing and annotation mechanisms. Supporting shared annotation in and around the score essentially embeds the creative process in the score, creating a rich and engaging multi-modal experience. However, the persistent nature of annotations in Daisy* may reduce their utility in longer-term collaborations—some form of erasing mechanism by users, or over time, may be necessary to reduce screen clutter and confusion. Also, Daisy* has primarily been used to support the process of shared composition, so it would be interesting to explore their use in more performative forms of interaction [28], where mutual engagement between participants is more improvisational than compositional.

Distributed music-making relies on task management, knowing who one's collaborators are and entertaining social interaction to glue it together. Daisy* provides elements of all of these, in addition to the ability to localize annotations using a simple shared graphical annotation mechanism. Future work will explore the role of persistence of the annotations and the score itself, as well as richer and more performative music-making experiences.

Acknowledgment

This research was supported by EPSRC grants EP/H042865/1, EP/E045235/1, GR/S81414/01.

References and Notes

1. G. Weinberg, "Local Performance Networks: Musical Interdependency through Gestures and Controllers," *Organised Sound* **10**, No. 3, 255–265 (December 2005).

2. J.T. Titon, *World of Music* (New York: Schirmer Books, 1996).

3. J.-B. Thiebaut, P.G.T. Healey and N. Bryan-Kinns, "Drawing Electroacoustic Music," in *Proceedings of International Computer Music Conference,* Belfast (ICMC 2008).

4. S. Nabavian and N. Bryan-Kinns, "Analyzing Group Creativity: A Distributed Cognitive Study of Joint Music Composition," in *Proceedings of Cognitive Science 2006* (2006) pp. 1856–1861.

5. See, for example, A.A. Sawchuk et al., "From Remote Media Immersion to Distributed Immersive Performance," in *Proceedings of ETP '03*, Berkeley, CA (2003).

6. G. Föllmer, "Lines of Net Music," Martin Watson, trans., *Contemporary Music Review* **24**, No. 6, 439–444 (2005).

7. T. Blaine and S. Fels, "Contexts of Collaborative Musical Experiences," in *Proceedings of the 2003 Conference on New Interfaces for Musical Expression* (NIME03), Montreal, Canada (2003) pp. 129–134.

8. G. Föllmer, "Electronic, Aesthetic and Social Factors in Net Music," *Organised Sound* **10**, No. 3, 185–192 (2005).

9. S. Jordà et al., "The reacTable," in *Proceedings of the ICMC2005,* Barcelona, Spain (2005).

10. G. Wang, "Designing Smule's iPhone Ocarina," in *Proceedings of the International Conference on New Interfaces for Musical Expression,* Pittsburgh (2009).

11. A. Barbosa, "Displaced Soundscapes: Computer-Supported Cooperative Work for Music Applications," Ph.D. Thesis, Universitat Pompeu Fabra, Spain, 2006.

12. Weinberg [1].

13. R. Mills, "Dislocated Sound: A Survey of Improvisation in Networked Audio Platforms," in *Proceedings of NIME 2010* (NIME10), Sydney, Australia (2010) pp. 186–191.

14. See S. Fels and F. Vogt, "Tooka: Explorations of Two Person Instruments," in *Proceedings of the 2002 Conference on New Instruments for Musical Expression* (NIME02), Dublin, Ireland (2002).

15. N. Bryan-Kinns, "Daisyphone: The Design and Impact of a Novel Environment for Remote Group Music Improvisation," in *Proceedings of DIS 2004*, Boston, U.S.A. (2004) pp. 135–144.

16. <Daisyphone.net>.

17. P. Marks, "Will Jamming Be the New Texting?" *New Scientist* **180**, No. 2418 (25 October 2003) p. 25.

18. Barbosa [11].

19. P. Burk, "Jammin' on the Web—A New Client/Server Architecture for Multi-User Musical Performance," presented at ICMC 2000.

20. See, for example, *Surface Tension*, Dana Centre, Science Museum, London, U.K., <www.danacentre.org.uk/events/2009/06/23/506> (23 June 2009).

21. N. Bryan-Kinns and F. Hamilton, "Identifying Mutual Engagement," *Behaviour & Information Technology,* DOI: 080/01449290903377103> (2009).

22. See M. Csikszentmihalyi, *Flow: The Psychology of Optimal Experience* (Harper Collins, 1991); R.K. Sawyer, *Group Creativity: Music, Theater, Collaboration* (Mahwah**,** NJ: Lawrence Erlbaum Associates, 2003).

23. N. Bryan-Kinns and P.G.T. Healey, "Decay in Collaborative Music Making," in *Proceedings of NIME 2006* (NIME06), Paris (2006) pp. 114–117.

24. C. Small, *Musicking: The Meanings of Performing and Listening (Music Culture)* (Wesleyan Univ. Press, 1998).

25. N. Bryan-Kinns and P.G.T. Healey, "Exploring Mutual Engagement in Creative Collaborations," in *Proceedings of Creativity and Cognition* 2007, Washington, U.S.A. (2007) pp. 223–232.

26. Bryan-Kinns et al., [15], [21], [25].

27. J.S. Olson et al., "Groupwork Close Up: A Comparison of the Group Design Process with and without a Simple Group Editor," *ACM Transactions on Information Systems* **11**, No. 4, 321–348 (1993).

28. J.G. Sheridan and N. Bryan-Kinns, "Designing for Performative Tangible Interaction," *International Journal of Arts and Technology,* Special Issue on Tangible and Embedded Interaction, **1**, Nos. 3–4, 288–308 (2008).

Manuscript received 1 January 2011.

Nick Bryan-Kinns is a senior lecturer in Graphical User Interface Design and leader of Interactional Sound and Music at the Centre for Digital Music, Queen Mary, University of London. His EPSRC-funded research focuses on exploring mutual engagement in creative collaborations, and he has been a panel member for the NSF CreativeIT program, chair of the ACM Creativity and Cognition 2009 conference and co-editor of special issues of journals such as CoDesign.