Interactional Sound and Music: Listening to CSCW, Sonification, and Sound Art

N. Bryan-Kinns
Queen Mary University of London
Interactional Sound and Music Group
Centre for Digital Music
London
E1 4NS
nickbk@dcs.qmul.ac.uk

R. Fencott, O. Metatla, S. Nabavian
Queen Mary University of London
Interactional Sound and Music Group
Centre for Digital Music
London
E1 4NS

J. G. Sheridan
BigDog Interactive Ltd
London

ABSTRACT
In this paper we outline the emerging field of Interactional Sound and Music which concerns itself with multi-person technologically mediated interactions primarily using audio. We present several examples of interactive systems in our group, and reflect on how they were designed and evaluated. Evaluation techniques for collective, performative, and task oriented activities are outlined and compared. We emphasise the importance of designing for awareness in these systems, and provide examples of different awareness mechanisms.

Keywords
Interactional, sound, music, mutual engagement, improvisation, composition, collaboration, awareness.

1. INTRODUCTION
Interactional[1] Sound and Music (ISM) involves multiple people interacting together using audio as the primary modality. Examples range from task oriented activities such as collaborative sonification, through to affective and aesthetic experiences such as interactive collective audio explorations. In this paper we present three dimensions which can be used to organise the requirements of such systems. We propose that at the core of designing for such systems is a focus on awareness. The key question then is: how to design awareness into multi-person audio experiences which have radically different design and evaluation criteria - evaluating awareness in a collective art experience is very different to evaluating awareness in collaborative sonification tools. As such, ISM is inherently multi-disciplinary, requiring an understanding of topics from managing collaborative work through to designing engaging artistic experiences.

2. ISM, NIME, and CSCW
Collaboration is fundamental to human creative endeavors, yet is under-explored in the design of interactive systems. This is especially important for the design of New Interfaces for Musical Expression (NIME) as music making and consumption usually involves more than one person. Indeed, [17] argue that the production and enjoyment of music is typically an open, collaborative, and 'ubiquitous'. In this paper we describe examples of ISM projects and reflect on how different theories and techniques have been used to inform the design of NIME systems in this area. We also reflect on the scope and appropriateness of these approaches, and highlight how NIME systems can in turn be used to inform the design of non-musical audio interaction. ISM and NIME are at the same time both broader and narrower in scope than each other. ISM is broader than NIME as it focuses on all uses of audio in collaboration, rather than primarily focusing on musical expression. Conversely, ISM is narrower in scope than NIME as it is concerned solely with multi-person interaction, whereas NIME does not make such a narrow definition. This overlap between the two areas provides useful methodological insights into each which we will explore in this paper. The key contribution that ISM can bring to NIME is the employment of CSCW frameworks to inform the design and evaluation of musical experiences.

[1] *Interactional*: capable of acting on or influencing each other

2.1 Features of ISM Activities
Our experience of designing Interactional Sound and Music systems (ISMs) over the last seven years has led to the identification of three key dimensions which can be used to distinguish ISM activities that people might get involved with. In contrast to NIME design dimensions for collaboration (e.g. [3]), our focus here is on the nature of the activities participants are involved in, i.e. the requirements on systems, rather than the features of systems. These dimensions are: **Situation** of participants i.e. whether they are co-located or remotely collaborating; **Focus** of the activity ranging from collaboration to collective action, i.e. whether the activity is task/ work oriented, or focuses on creative engagement between participants; **Immediacy** of the activity, that is, whether an action is immediately considered a product in itself, or whether the product is iteratively refined over time.

Designing for these activities fundamentally requires an understanding of awareness. Even if an activity is co-located and compositional, participants need to be able to co-ordinate their actions. In this paper we reflect on our explorations of designing both audio and visual awareness mechanisms for ISMs.

2.2 Awareness
The study of the technological requirements and implications of people working together has largely been conducted in the field of Computer Supported Collaborative Work (CSCW). For example, studies have examined shared document editors such as ShrEdit [8][14], group decision support e.g. [1], and collaborative brainstorming systems [12]. To date this work has

typically been work oriented, focusing on conventional office and text document based activities. However, it has led to the characterisation of design features for collaboration such as awareness mechanisms e.g. [8][9], and the importance of shared representations e.g. [15]. Whilst shared representations may be crucial to some forms of ISM (cf. [4]), we argue that, from our experience, awareness mechanisms are crucial to all ISM systems. Moreover, studies of joint music composition [13] and free improvisation [10] have repeatedly shown the importance of mutual awareness of actions in music making.

3. From Collaboration to Collective Action

The evaluation of ISM systems is driven primarily by the kind of activity that participants are involved in i.e. the *focus*. In this section we outline how evaluating task focused activities is quite different from evaluating activities focused on engagement. As we shall see, for task oriented evaluations we can draw directly on HCI and CSCW evaluation techniques such as assessment of task completeness, efficiency, participant comprehension, and so on. For activities that are not task focused, but rather focus on the engagement between participants (e.g. in a group improvisation) we employ the concept of mutual engagement, and outline some techniques for assessing engagement between participants. Our position is that mutual engagement occurs when people creatively spark together and enter a state of group flow [6]. The distinguishing characteristic of mutual engagement is: “it involves engagement with both the products of an activity and with the others who are contributing to those products” [ibid]. Of course, people collaborating in task oriented activities may (hopefully) become mutually engaged, but entering into a state of group flow is not the primary focus of the activity.

3.1 Remote Collective action

We have developed a number of iterations of the Daisyphone system [4] which supports remote music composition of short loops. In this system participants collaboratively edit a short (1 minute) loop of music. Figure 1 illustrates the user interface - collaborating users each see the same user interface. Coloured shapes indicate different peoples’ notes added to the loop. The grey line rotates around the circle, and as it does so, music is created from the notes it passes over. Each person can add different kinds of sound, and can edit their own as well as others’ contributions. They can also add free-style annotations over the interface.

We have conducted several lab based and longitudinal web-based studies of the effect different awareness mechanisms, and

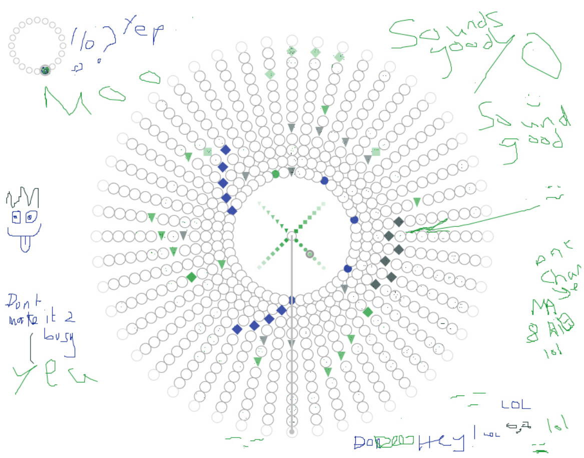

Figure 1. Daisyphone in use

annotation mechanisms [6]. Most recently we have shown that providing shared visual annotation mechanisms and visual awareness of the identity of co-composers significantly increases mutually engagement in interaction [6].

Through our studies we iteratively refined a set of measures of mutual engagement based on participants’ interaction, and a robust questionnaire which can be used to compare different interfaces. The Mutual Engagement Questionnaire (MEQ) is used to compare two or more user interfaces. Participants use all versions of the interface, and then indicate which statements on the MEQ were most suited to which interface. This provides a clear, and usually statistically reliable measure of which interface was most mutually engaging. The statements in the MEQ are grouped into four categories (these category titles are not conveyed to the participants): **Satisfaction** with the product; **Feelings** of enjoyment or *flow* cf. [7]; Sense of **collaboration**; **Usability**.

We also developed a set of reliable measures of mutual engagement based on participants’ actions which we refer to as *interactional profiling*. These are: number of contributions; amount of co-editing (i.e. editing each others’ contributions); evidence of convergence of musical ideas (i.e. alignment and repetition of musical motifs).

We have been conducting studies of a new version of Daisyphone – Daisyfield – which allows multiple loops to be edited at once, and allows spatial layout of contributions. Again, colours represent who contributed what. In Daisyfield we can additionally display the current mouse pointer position of other users - these are represented by small circles and provide some awareness of others’ focus of activity. The **key challenge** here is: how to support awareness at a distance which does not overload participants.

3.2 Co-located Interaction

With co-located interaction the design question is more about how to augment participants’ existing natural awareness of each other by virtue of their physical proximity to encourage more mutually engaging experiences.

3.2.1 Exploratory

In contrast to music making, we have developed systems which focus on using sound to explore other phenomena. One of our systems, Sensory Threads [5], is a multi-person mobile experience in which participants sense the imperceptible around them through a responsive real-time soundscape. Participants are sent on a group expedition (four people), where each person wears a device which records data from one type of sensor (heart rate, light, sound, spatial density) as illustrated in figure 2. Each person’s individual sensor stream is sent to a ‘heart’ computer carried by one of the explorers. The four sensor streams are then used as input to an interactive soundscape which is produced by the heart computer. Participants listen to the soundcape using wireless headphones as they move around the city. The heart computer also records the location of the explorers using GPS which allows the system to track of where the explorers have been. Each sensor influences a different aspect of the generated soundscape. This creates a complex group dynamic, where participants engage with their own actions in the space as well as the communal sensing experience. There is no iteration of the participants’ contributions – sounds immediately happen and then disappear, the interaction is of-the-moment. In this situation we aim to design awareness of each participants’ sensor readings, and hence the collective state of the exploration. The experience is radically different to that of collective music making with

Figure 2. Sensory Threads

Daisyphone as participants only use the auditory modality, and the experience is not about making music per se, but rather exploring an environment through music. Sensory Threads is also different to other pervasive games such as Uncle Roy All Around [2] as Sensory Threads is not competitive, and the primary modality is sound.

The purely audio nature of Sensory Threads leads to a number of additional awareness mechanism design questions. First, as we had no recourse to visual media to convey identity, we used different sounds and modulations of sounds for each data stream to convey contributors' identity. We drew on sonification research [11] to inform the sound design process. In particular, we used mappings of sensor data to: pitch, pulse, filter, and density of sound. It would have been tempting to map a noise sensor to the volume of static noise in the soundscape, but whilst this is a literal mapping, it may be confusing as the louder the environment becomes, the louder the static noise in the soundscape would be. Instead, we selected sounds and modulations that somehow captured the essence of the phenomena being sensed. The **key question** here is: how to provide awareness and interaction using **only audio**.

3.2.2 Performative

One of the great benefits of audio based interaction is that it has the potential to draw people into new forms of collective action. We developed a system called uPoi [16] which specifically seeks to draw bystanders into collective action and become not just participants in the group activity, but hone their skills to become performers. With uPoi people swing two computational augmented poi (balls on strings) around their body. Data about the movement is wirelessly transmitted to a base computer which then generates sound and visuals based on the action of all participants. Usually up to four people interact at once, giving us eight sensor readings in real-time. In keeping with Sensory Threads, we assign a specific style and modulation of sound to each participant to create a sense of auditory awareness of each others' actions. Again, the system concerns itself with immediate production of sound, not with editing of the product in any way. In contrast to Sensory Threads, we also project a visual representation of the interaction into the collective performance space. In keeping with Daisyphone, the visual imagery is colour coded to provide a visual awareness of who is contributing what to the emerging audio visual experience. As with Daisyphone, we found that providing no visual cues to identity reduced people's mutual engagement. We have used the Performance-Triad model [16] to understand how people engage with each other through and around uPoi. In particular, we have identified transitions in people's behaviour from observing, to participating, and on to performing where they convey emotions through uPoi. Performance is typically indicated by complex technical skills and interaction with other uPoiers, for example by playing sounds off each other, or pausing to create dramatic effect for other performers.

3.2.3 Improvisational

We have developed Daisyphone into a co-located improvisation system called Daisychain. In Daisychain co-located participants interact with handheld devices to collectively create music. As with Daisyphone, a short loop is at the core of the interaction. However, with Daisychain, the loop is spatially distributed around the participants – notes hop from one device to the next and can be edited by any participant. This makes Daisychain unlike other co-located interactive music making such as BeatBugs [18] which focuses on sharing beats and musical motifs. Again, we provide a shared representation of the loop. This interaction makes Daisychain somewhere between collective composition and improvisation - there is scope for immediate improvisation as well as iterative refinement and editing of contributions. Furthermore, in Daisychain we introduce a personal improvisation space which allows participants to try out sounds on their own device before sending them into the loop. We have also introduced a controllable decay which moves the activity from purely editing to somewhere between editing and improvisation.

The **key question** here is: how can the **spatial arrangements** of participants be used to reinforce awareness of collective action.

3.2.4 Collective composition

Informed by analyses of Joint Music Composition [13] we have been developing co-located collective music making tools. These provide richer musical expression and composition than Daisyphone or Dasiyfield, and tend towards being task oriented, with the aim of the collective interaction being to create a good piece of music. We are currently undertaking studies of manipulating participants' control of levels of privacy by providing different levels of access to each others' contributions. It may be that for more task oriented activities it is important to provide some privacy of action, and some form of access control over one's own products. Again, we have built audio and visual awareness mechanisms into the design to support collective composition, and are assessing the system through the MEQ and interactional profiling. Early results suggest that the provision of private interaction spaces increases participants' mutual engagement.

The **key question** here is: what forms of **privacy and awareness** provide the most mutually engaging experience.

3.2.5 Problem solving

Audio is not just for music making. We have been exploring the design and evaluation of audio only collaborative diagram editing tools - a form of interactive sonification [13]. We use the term collaborative here to distinguish these activities from the less task focused, collective, activities of group music making, and performative expression. We are now undertaking studies of collaborative audio only editing of UML diagrams. At the system's heart are, as would be expected, awareness mechanisms. As with Sensory Threads, we have developed a system which is purely audio. So, again, we have carefully designed the audio awareness mechanisms by employing complimentary sounds and timbres for each collaborator which are distinct, yet aesthetically appealing. Task oriented systems such as this can learn a lot from more artistically oriented systems such as Sensory Threads, especially with reference to creating and using sounds which are complimentary, intuitive, and do not become annoying over extended periods of use. As the system is intended for task focused problem solving, we will not be using the MEQ or Performance-Triad, but instead will be evaluating the system using more conventional HCI based measures such as efficiency of interaction, quality of end product, and participants' comprehension of the shared activity. It will be interesting to compare task based and non-task based evaluations of the use of audio for awareness in audio only interaction.

The **key question** here is: how to design **aesthetically pleasing** awareness mechanisms which clearly convey the underlying activities and data.

4. Summary

We presented a number of examples of Interactional Sound and Music systems. What is key here is the fundamental importance of providing awareness mechanisms to support the group activities. The focus of the activities, whether they are co-located or not, and whether audio is used alone or in conjunction with other modalities, all contribute to informing the design and evaluation of our awareness mechanisms. In this paper we summarized a number of design and evaluation techniques and considerations which can help in future research.

5. ACKNOWLEDGMENTS

Sensory Threads was developed by Demetrios Airantzis, George Roussos, Jenson Taylor (Birkbeck, University of London); Frederik Lesage (London School of Economics and Political Science); Joe Marshall (University of Nottingham); Alice Angus, Giles Lane, Orlagh Woods (Proboscis); Lorraine Warren (University of Southampton); Nick Bryan-Kinns, Robin Fencott (Queen Mary University of London). Development work took place between September 2008 and June 2009.

Sensory Threads is funded by the EPSRC Creator research cluster (EP/G002088/1: New Research Processes and Business Models for the Creative Industries), and the EPSRC Platform Grant: Centre for Digital Music (EP/E045235/1).

Robin Fenott and Shahin Nabavian are funded by the EPSRC.

Daisyphone was initially funded by EPSRC grant GR/S81414/01.

6. REFERENCES

[1] Applegate, L., Konsynski, B. R., and Nunamaker, J. F. (1986). A group decision support system for idea generation and issue analysis in organization planning. In *Proceedings of ACM CSCW'86*, 16-34.

[2] Benford, S., Crabtree, A., Reeves, S., Flintham, M., Drozd, A., Sheridan, J. G., and Dix, A. (2006). The frame of the game: Blurring the boundary between fiction and reality in mobile experiences. In *Proceedings of ACM CHI 2006*, 427-436.

[3] Blaine, T. and Fels, S. (2003). Contexts of collaborative musical experiences. In *Proceedings of NIME 2003*, 129–134.

[4] Bryan-Kinns, N. (2004). Daisyphone: The Design and Impact of a Novel Environment for Remote Group Music Improvisation. In *Proceedings of ACM DIS 2004*, Boston, USA, 135-144.

[5] Bryan-Kinns, N., Airantzis, D., Angus, A., Fencott, R., Lane, G,, Lesage, F., Marshall, J., Martin, K., Roussos, G., Taylor, J., Warren, L., & Woods, O. (2009). Sensory Threads: Perceiving the Imperceptible. In *Proceedings of 5th International Conference on Intelligent Environments* (IE'09).

[6] Bryan-Kinns, N., & Hamilton, F. (2009). Identifying Mutual Engagement. *Behaviour & Information Technology*.

[7] Csikszentmihalyi, M. (1991). *Flow: The Psychology of Optimal Experience*. Harper Collins.

[8] Dourish, P., and Bellotti, V. Awareness and Coordination in Shared Workspaces. In *Proceedings of ACM CSCW'92*, 107-114.

[9] Gutwin, C., and Greenberg, S. (2002). A Descriptive Framework of Workspace Awareness for Real-Time Groupware. *Computer Supported Cooperative Work*, Vol. 11, 411-446.

[10] Healey, P. G. T., Leach J., and Bryan-Kinns, N. (2005). Inter-play: Understanding group music improvisation as a form of everyday interaction. In *Proceedings of Less is More — Simple Computing in an Age of Complexity*. Microsoft Research Cambridge

[11] Hermann, T. (2008). Taxonomy and Definitions for Sonification and Auditory Display. In *Proceedings ICAD 2008*, Paris, France.

[12] Hymes, C. M. and Olson, G. M. (1992). Unblocking brainstorming through the use of a simple group editor. In *Proceedings of ACM CSCW'92*, 99-106.

[13] Metatla, O., Bryan-Kinns, N., & Stockman, T. (2008) Constructing Relational Diagrams in Audio: The Multiple Perspective Hierarchical Approach. In *Proceedings of ASSETS2008*, Halifax, Canada.

[14] Nabavian, S., and Bryan-Kinns, N. (2006). Analysing Group Creativity: A Distributed Cognitive Study of Joint Music Composition. In *Proceedings of Cognitive Science*, 1856-1861.

[15] Olson, J. S., Olson, G. M., Storrøsten, M., and Carter, M. (1993). Groupwork close up: a comparison of the group design process with and without a simple group editor. *ACM Transactions on Information Systems*, Vol. 11, No. 4, 321-348.

[16] Robertson, T. (1997). Cooperative Work and Lived Cognition: A Taxonomy of Embodied Actions. In *Proceedings of ECSCW '97*, 205-220.

[17] Sheridan, J. G. and Bryan-Kinns, N. (2008). Designing for Performative Tangible Interaction. *International Journal of Arts and Technology. Special Issue on Tangible and Embedded Interaction*.

[18] Sloboda, J. A. and O'Neill, S. 2001. Emotions in Everyday Listening to Music. *Music and Emotion: Theory and Research*. Oxford University Press, UK, 415–430.

[19] Weinberg, G., Aimi, R., and Jennings, K. (2002). The beatbug network: a rhythmic system for interdependent group collaboration. In *Proceedings of NIME 2002*