UNDERSTANDING AURAL FLUENCY IN AUDITORY DISPLAY DESIGN FOR AMBIENT INTELLIGENT ENVIRONMENTS

Milena Droumeva

Faculty of Education
Simon Fraser University
mvdroume@sfu.ca

Ron Wakkary

Associate Professor
School for Interactive Arts and Technologies
Simon Fraser University
rwakkary@sfu.ca

ABSTRACT

This paper presents the design and some evaluation results from the auditory display model of an ambient intelligent game named socio-ec(h)o. socio-ec(h)o is played physically by a team of four, and displays information via a responsive environment of light and sound. Based on a study of 56 participants involving both qualitative and preliminary quantitative analysis, we present our findings to date as they relate to the auditory display model, future directions and implications. Based on our design and evaluation experience we begin building a theoretical understanding for the unique requirements of informative sonic displays in ambient intelligent and ubiquitous computing systems. We develop and discuss the emerging research concept of *aural fluency* in ambient intelligent settings.

1. INTRODUCTION

socio-ec(h)o is a prototype environment for group play whose goal is to explore issues of design, use and evaluation for ambient intelligent systems. Ambient intelligence (AmI) computing is the embedding of computer technologies and sensors in architectural environments that combined with artificial intelligence and multi-modal displays, respond to and reason about human actions and behaviors within the environment. Our broader research with socio-ec(h)o is focused on several related issues including motion sensing, light and sound display design, user modeling, and interaction models. The overall research goal is to understand how to support groups of participants as they learn to manipulate an ambient intelligent space, as well as to better understand how to design ambient components of a responsive environment capable of providing this type of support. The research questions are numerous in a project of this nature and yet immersive and embodied interaction does not lend itself to reducible and isolatable variables that can be measured independently. Given this, our study design has focused on an ecological investigation and theory-building, aiming to provide broad, yet particular set of heuristics that help describe and make sense of the different system and display components of AmI environments. The empirical quantitative results, which include a survey instrument, time of completion and accuracy measures, as well as video coding analysis, are only one part of the picture, providing supportive evidentiary data, which, along with our overall observations and design process reflections, contribute to building a fuller picture of the role, unique characteristics and requirements of ambient intelligent systems.

This paper will specifically address the auditory display design of socio-ec(h)o. First, we make room for a methodological discussion related to approaching rich, complex settings in order to situate our inquiry within the larger paradigm of interaction design as well as ecological methods of auditory perception research. Based on the particular model of sonic feedback in socio-ec(h)o, we offer some results to date along with a preliminary attempt to formulate key factors and concepts towards building a framework of auditory displays in ambient intelligent environments. A second reason for a focus on the methodological aspects of the study is that the main contribution of the paper, a concept of *aural fluency* arose from an emergent set of questions afforded by this approach to research. In this paper, we outline the initial definitions and resulting heuristics for *aural fluency* that we believe will lead to further investigation into auditory display design for ubiquitous computing systems.

Figure 1. *socio-ec(h)o gameplay, level 4.Participants work together to solve the puzzle Big Bang! and build a big fire from quiet crackles.*

2. METHODOLOGICAL DISCUSSION

There is the design challenge of crafting an informative auditory display for an ambient intelligent system, and then there is the research challenge of building knowledge about the role of sound in ubiquitous computing environments, as well as the more concrete goal of adding to auditory perception research via ecological investigations. Since at present there are still relatively few examples of complex auditory displays where sound is made meaningful, essential, communicative, engaging and social within a ubiquitous computing setting, it is hard to draw methodological and epistemological examples to guide our present inquiry. In addition, the research into actual ubiquitous computing environments is still in the process of bridging knowledge between psychology, cognitive science, human-computer interaction. Controlled experiments into auditory display issues are simply insufficient at this point in time in addressing these complex requirements of ubiquitous computing, as we don't yet understand enough about their foundational components. The issues that motivate this current investigation have to do with both understanding the role and requirements of informative sound feedback in ambient intelligent systems, as well as determining best practices in perceivability [1] of sound in such contexts.

In traditional psychology experiments on auditory perception, as well as in quantitative human-computer interaction methods, there is rigid stimulus-response formula. Such studies are good for determining absolute, static thresholds and values of human auditory perception. Some auditory display studies, particularly Kramer and Gaver's work [2] in auditory icons, and more currently the works of Kramer, Walker, Brewster and Barrass in earcons and sonification [3, 4], have gone farther by situating the stimulus-response formula in a real task situation with a meaningful activity that uses sound feedback. Among others, McAdams' and Ballas' work on environmental sound perception [5] also takes the inquiry further by examining how people's experiences with everyday, environmental sounds happen within different models of perception, cognition and action, than is suggested by laboratory tests with synthetic tones. Sonification studies, as a subset of auditory display research, focus on determining measures and thresholds of auditory perception responses of continuous, changing sound in terms of scaling, polarity, data-to-parameter mapping and spatialization [1, 3]. Contemporary sonification studies take context and other user-centered design issues into account in the research inquiry, as well as emphasize considerations such as time-prolonged exposure to sonifications; auditory and cognitive fatigue; and aesthetic aspects as an element of efficiency of the sonification [6]. The use of non-traditional sound categories for display - using vocal or other non-synthetic content [7, 8] as well as the role of interaction sonification [8] have also been investigated, though by few. In any case, the unit of analysis is always the human response, measured as a threshold, a just noticeable difference, a correct or incorrect response – all in an attempt to capture generalizeable human perceptual and cognitive abilities and predict reactions.

On the other end, while the fields of human-computer interaction and interaction design place emphasis on user-centered design in developing and researching technology, they rarely deal with investigating sound in a significant, systematic manner. When they do, it is often through traditional evaluation and usability methods, which still may not provide a setting that is ecological and holistic enough to allow for building and validating contextual, situated knowledge about sound in ubiquitous computing environments. As Gibson points out, "…awareness is rooted in meaningful experience of the environment: thus ecological validity results from studying subjects/people in their own natural environment, in motion, in active exploration. For people this environment is social, cultural, systemic, economic, political, etc." [9]. Examining social contexts, group interactions, and embodied experiences and their interplay with sound, listening and soundscapes, could not therefore be achieved through traditional psychology or usability methods alone, and requires, we argue, adopting the situated, activity-based approach of participatory design methods in order to achieve this task better [10, 11, 12].

3. PROTOTYPE DESCRIPTION AND AUDITORY DISPLAY MODEL

The socio-ec(h)o game is played by a team of four players and features six levels of increasingly difficult word puzzles solved by coordinated body positions and movements – see Table 2 for puzzle titles – for example, in Lo and Behold all four players had to be crouching low in the middle of the space and hold that position for around 3 seconds. For more details on the game prototype see [13]. The environment is responsive to players' actions through abstract light and sound and the ambient response is driven by an artificial intelligence module that reasons about players' actions in space. Players' movements are tracked using a Vicon motion capture system (www.vicon.com). The goal was to create an ambient system that enables users to learn, use and manipulate it for a problem-solving activity. Based on our preliminary participatory design workshop results, already outlined in [14], we established that the interface required a gradient rather than direct response in order to best represent game states, as well as anticipate and direct players' actions. In keeping with AmI design we wanted to move away from confirmatory "sound/no sound" feedback, which only signals if users do something right. Moreover we wanted to move beyond discreet incremental changes that still promote an "action-reaction" model of communication and are known to pose problems to short-term auditory memory [15]. Instead, we wanted to create a dynamically gliding intensity of sound that not simply responds but directs user actions and relies on perception of change and just-noticeable differences of different sound parameters. We even introduced an artificial 2-second delay in updating sonic feedback to the game state, in order to push the interaction away from an instant-feedback scenario and towards more complex, attentive and reflective listening – acting model similar to a form of data sonification.

We termed this approach *intensity-based gradient sound feedback model* [more in 16]. This intensity gradient shifts smoothly between layers of sound and provides not only narrative complexity but also relays information and guides player actions through ambient real-time sonification (see Tables 1 and 2) - that is, tells players how 'close' they are to solving the puzzle. Naturally, principles of data sonification came at the basis of the auditory display model. The principles we borrowed included the concepts of scaling, polarity, data-to-parameter mapping and spatialization [3, 17, 18, 19, 20]. In addition, we adhered to general principles of psychoacoustics for just noticeable differences, effectiveness of different parameter shifts (pitch, versus amplitude, tempo or timbre, as an example) [19, 21]. In addition to guidelines from sonification frameworks [1, 3], our sonic feedback borrowed principles from soundscape composition and everyday listening [2, 22]. Finally, our choice of activity format – a game – came because games are suitable for research, particularly one with an ecological onus, since they provide avenues of natural engagement, while at the same time game mechanics allow careful constraints to be built into the experience, and the quantifiable outcome (game solution) allows for testing basic effectiveness, comparing times of completion, and capturing rich interactions. The way sound display fit into the game as a core mechanic was in following the natural skill mastery progression of the game – the soundscape in Level 1 was perceptually easy to interpret (see Table 1) while the soundscape for Level 4 was significantly more ambiguous. Thus ambiguity, perceptual and aesthetic, rather than being avoided, became a core part of the game experience itself.

3.1. Audio Display Schema

The audio display for socio-ec(h)o consists of three components: a real-time ambient sonification engine that has a different soundscape for each level of the game; an anticipatory feedback sound to signal when all participants are working together towards a goal; and a confirmatory feedback sound, which signals the completion of the goal and the progress to the next level (see [16] for details). The latter two feedback signals are consistent through all levels, so as to create a coherent interpretation of success. For example, an answer to the word puzzle “lo and behold” is for all players to crouch low and be still. Further, they must hold this configuration for a short period of time, typically 4 seconds. This ensures that the actions were purposeful and knowingly completed rather than an accidental formation. As they crouch lower, the sound feedback intensifies, and if they stand up again the ambient sound will de-intensify. When they achieve the desired position and hold it for a few seconds, an anticipatory feedback sound occurs signaling to the players that they’ve done something right and if they keep at it they will complete the solution. Upon sustaining the configuration, the level changes with an audio reward and a short video reward.

Levels	Core soundscape	Approaches to intensity
L1	Soft abstract musical soundscape	Amplitude only
L2	Soft abstract musical soundscape	Phaser-tempo (slight pitch and amplitude adjustments)
L3	Clinking pebbles soundscape	Pitch only (slight amplitude adjustments
L4	Fire sounds	Cross-fade between five recorded fire soundscapes of increasing intensity
L5	Water stream sound	Phaser-tempo (slight pitch and amplitude adjustments)
L6	Forest ambience	Cross-fade between five recorded forest soundscapes from calm day to thunder storm

Table 1: Detail of the soundscapes and approaches to displaying auditory intensity for each game level.

As demonstrated in Table 1, sonification principles were used in conjunction with one another and alongside more narrative-based intensity approaches. Specifically, the cross-fader between five different recorded soundscapes of increasing intensity was something we wanted to test out against more aurally and perceptually simple approaches to intensity such as sliding pitch or rising/falling amplitude. Further, the introduction of this narrative sonic component supplemented the game mechanic of rising gameplay challenge – becoming harder to interpret.

Several factors posed challenges to our exploration of the effectiveness of these auditory design approaches. The fact that each level has a different puzzle presents a subjective element of experience that comes in addition to each level using a different soundscape and a different intensity model. Table 2 outlines some of the specific scaling values used in the Max/MSP audio engine. Even though sound parameter changes are stacked to complement each other, the scaling values of main variables were kept similar as much as possible with different base soundscapes. In addition, amplitude adjustments were necessary, especially in the negative polarity cases to counteract effects of pitch and timbre for the equal loudness effect. Negative polarity was chosen in the tempo cases to support a narrative and embodied connection to the puzzles – in the case of Level 2 the participants had to slow down in order to solve the puzzle, and at the same time the sound’s pulsing tempo was going down instead of up. Similarly, in Level 5 the soundscape was slowing down to encourage participants to experiment with still and slow movement, which was again required in order to achieve the desired configuration. Alternatively, in Level 3 users had to run faster than normal and the positive polarity of pitch and upbeat, crisp timbre of crunching pebbles provided a good connection between activity and response. Lastly, in the case of Level 4, the fire soundscape, which built up from crackles and sizzles to a full-blown bonfire supported the motif of the puzzle, which was Big Bang! The physical solution was to start in the middle and ‘make the fire grow’ then move outwards to the edges standing up.

Levels	Word Puzzle	Polarity	Exact Scaling Values
L1	Lo and Behold	Positive	20 –140 (max 157)
L2	Sloe and Low like a Plum Turtle	Negative	8.00 – 0.00 phase cycle (0.00 is full phase) 1.20 - 0.8 pitch (1.00 is normal pitch)
L3	All Rolling in a Bowl	Positive	0.7 – 1.9 pitch bend Amplitude adjustment
L4	Big Bang!	Positive	Cross-fade fire Amplitude adjustment
L5	Gazing over Waves	Negative	8.00 – 0.00 phase cycle (0.00 is full phase) 1.4 - 0.8 pitch (1.00 is normal pitch)
L6	Swinging in the Ring of Fire	Positive	Cross-fade forest ambience Amplitude adjustment

Table 2: Detail of the game puzzles, polarity and scaling for each game level.

4. STUDY DESIGN AND RESULTS

In order to evaluate the effectiveness of our prototype design, as well as to investigate and build knowledge about interactions in space, ambient intelligence and auditory display for ubiquitous computing, we created a study protocol based on playing the socio-ec(h)o game. As mentioned in the beginning, a project of this magnitude entails many research questions and points of interest. The study design was meant to address a lot of them, however, not all are relevant to our current discussion of auditory display design and auditory perception, and thus would not be discussed here. 14 teams of 4 were enlisted to play the socio-ec(h) o game. The experience was conducted in two parts – in the first participants have to complete the first four levels within a maximum time of 60 minutes. In the second part participants have to attempt to complete the remaining two levels within 15 minutes. The research questions we had upon entering the evaluation stage of this project were as follows: 1) How informative was the ambient sonic feedback within our responsive environments? 2) Did our approach to designing the auditory display support participants progression in the game and to what degree? 3) Is there any significance in the effectiveness of different soundscapes and approaches to conveying information that we encoded?

4.1. Research Instruments

All study sessions were videotaped for the purpose of video analysis later. The artificial intelligence module also generated a log of participant activity displaying game success, as well as position in space, as outputted by the motion capture system. The same intensity measure drives the environmental response, so by looking at the log we know or can recreate the precise auditory display at any point of the experience for later analysis. Time of completion measures were collected, along with verbal transcripts, video analysis and finally, after the game portion, participants were given a survey to complete. The survey instrument contained a combination of Likert-scale questions and open-ended questions relating to the effectiveness of the system's response in helping, and guiding problem-solving, and creating an experience for them. Due to the enormity of the empirical data on this project, the numerous research questions, as well as lack of clear models to follow when interpreting results from complex, ecological studies of ambient intelligent systems we are still in the process of analyzing and interpreting results. To date, we have the survey results, time of completion measures, and some preliminary video analysis. In this paper we'll focus specifically on the survey results, both Likert-scale answers and more open-ended comments. In keeping with design methods process we also investigate emerging research questions and offer conceptual leads and future directions.

5. FINDINGS AND DISCUSSION

At a base level, our evaluation study of socio-ec(h)o aimed at testing the effectiveness of our design as demonstrated by times of completion (ToCs) and successful solution measures. In terms of the auditory display, we hoped to see some consistency in performance in certain levels across teams that could point to a design success or flaw of a particular approach to sound feedback or intensity. The survey and transcripts in turn were meant to serve as indicators of what participants perceived and thought about the system.

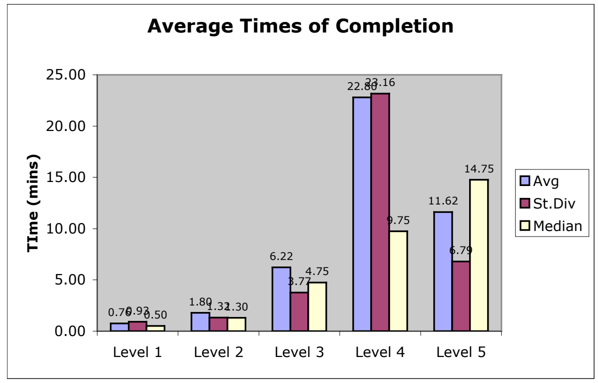

Figure 2. *Average times of completion, standard deviation and median values for Levels 1.2.3, 4, and 5.*

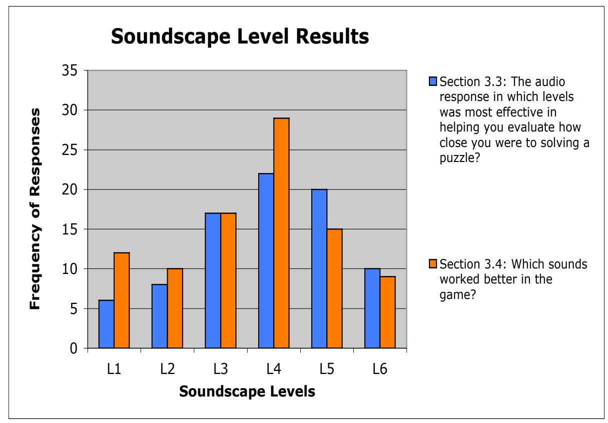

Figure 3. *Breakdown of results of how effective each level was in both its directive qualities and overall like.*

Q #	Question Wording	Avg	Medi	StDv
Q 1.2	1.2. Did the system require a large effort to learn?	3.07	3.00	1.07
Q 2.2	2.2. How well did the system provide useful cues in helping you evaluate how close you were to solving a puzzle?	3.58	3.75	0.46
Q 2.3	2.3. How well did the system respond when you were having problems solving a puzzle?	2.9	2.25	0.59
Q 3.1	3.1. Do the lights and audio work well together to help you in the game?	3.54	3.75	0.67
Q 3.2	3.2. How effective was the audio feedback in helping you solve the puzzles?	3.92	4.5	0.68
Q 3.5	3.5. How well does the virtual system (lights, audio, video) integrate with the physical and architectural space?	3.52	4	0.52

Table 3: Detail of survey question wording and results on a Likert scale out of 5.

In fact what we see in the results is no consistency at all in ToCs of different teams. Not only did ToCs vary greatly across levels (see Figure 2,) but they also varied greatly within a given level across different teams. As the most striking example, Level 4 demonstrates a range of completion times between half a minute and 53 minutes! What is curious and cannot be explained by these performance results are participants' ratings of the effectiveness of soundscapes and experience with audio response in different levels. In contrast to the completion times the questionnaire results are evenly distributed (see standard deviation values in Figure 3) and overall participants felt quite positive about the system's accuracy and effectiveness regardless of their individual performance. Figure 3 shows most participants both preferred the soundscape of level 4 and thought it worked best, thus showing no correlation between performance and preference for sonic feedback. So why did some teams take a long time to complete the puzzles, if they thought the sound response was good anyway and the game experience was positive?

To begin to address this question we analyzed the open-ended survey questions as well as the verbal transcripts from the sessions. While occasional players did not find the sound feedback useful, the majority of participants reported the sound being helpful and many described how 'the warmer sound told us we were close' while 'the negative sound meant we were doing the wrong thing'. The majority also reported the sound being more informative than the lights system (which incidentally was not designed as an intensity gradient, but used a pre-determined schema of light cues). The main suggestion for improvement on the sonic feedback was to make it more instant – many players perceived a lag, which they felt was confusing. The second main suggestion was for the feedback to be clearer in its representation of intensity – perhaps utilizing a negative intensity, signifying when players are definitely on the 'wrong' track.

Investigating the relationship between soundscape and intensity approach on one hand, and puzzle theme and physical configuration on another, revealed interesting results. Several participants commented in-game on Level 4's soundscape being 'the sound' of *Big Bang!* and related their progress in the solution to building up the fire and the concept of explosion with the word puzzle *Big Bang!* This aspect could in turn explain the popularity of Level 4's soundscape in the survey results (see Figure 3). Alternatively, many participants mentioned being confused or frustrated with the sound intensity approach in level 5, which uses a negative polarity (water pulse gets slower and slower as participants progress in the game). They specifically commented on the sound's change not being 'like waves' – referring to the word puzzle *Gazing over Waves*. Finally, a few participants mentioned that they liked level 3's sound because it 'sounded exactly like what they were supposed to do to solve it'.

These comments, along with the survey results and transcripts begin to reveal the role of sound feedback in socio-ec(h)o as one that is interconnected with notions of embodiment and multi-modal interaction and communication. While the multitude of variables in this experience makes it hard to form solid connections between sound feedback and performance, the ensuing discrepancies help re-frame and focus our analysis. Re-framing and re-problematizing an existing investigation is a common and necessary part of design research, and is particularly valuable in this less explored area of auditory display design for ubiquitous computing. Thus in addition to looking for salient and consistent results, we also looked for new angles of interpretation for the study results.

The main motivator for our emergent research questions is the unanswered question of why performance varied considerably while attitudes towards the system and the auditory display were positive but more importantly consistent. A careful look into the questionnaire results revealed a possible relationship between progression and skill acquisition, which led us to consider the concept of *aural fluency* as a skill in its own right. In other words, participants developed a certain competency about interpreting the sonic feedback as they played. Groups who were less 'successful', i.e. took a long time to solve a puzzle had the opportunity to build up an aural fluency overtime which allowed them to interpret the system's feedback better, with more subtlety, and resulted in a deeper connection to the system and more positive view of the auditory display as helpful – exemplified by their positive survey results. The concept of *aural fluency* may not directly explain how better performance can be supported by ambient audio display however it does suggest a "necessary but not sufficient condition" for that support. Interactors must *first* establish a level of fluency in understanding and responding to the dynamics of subtle, complex and changing audio, not unlike acquiring a new language before achieving gains in performance or higher levels of communication.

5.1. Exploring Aural Fluency

In line with design research, we take the newly problematized concept of *aural fluency* and first examine it through the lens of existing literature, then look for specific occurrences and pieces of evidence of it in our study results. From auditory perception research, the problem of aural training is well known. In the earcon and auditory icon fields a user must often formally or informally learn abstract sound signals and their association to digital or real actions [15, 17, 20]. Thus concept-to-signal mapping and especially built-in listener training is an essential part of designing auditory display systems. In data sonification, short-term auditory memory poses problems to users' ability to correctly and precisely interpret data values and trends from a continuous flow of sound. This problem has been explored in a few studies, by introducing auditory scaffolding such a reference tone [19] and contextual information [20] as well as through interactive sonification [8]. None of these studies however address the long-term effects of use on aural competency when interacting with a technological system. Further, the type of auditory literacy and fluency that is needed and seems to develop in more physical, situated technological environments such as ambient intelligent spaces seems more akin to everyday listening, as articulated by Truax and Schafer (in natural settings), as well as Gaver (in more technological settings) [2, 22]. It is situated in a social, contextual and shared physical environment and takes on characteristics of listening in everyday life, including dynamic shifts of attention, listening comprehension and aural expertise. As Jonas Löwgren writes in a new paper on fluency in augmented spaces, " the nature of our design material [of interaction design] is information and communication technologies, and unlike the products of industrial design and traditional architecture, it is temporal, as well as spatial" [23]. Aural fluency, we propose, is not simply another way of introducing auditory training issues, it constitutes a tangible connection between auditory perception, narrative and embodiment – that is, users build up aural fluency as a holistic experience of entails specific type of sound feedback, related to a specific set of movements and a particular narrative theme. So if aural fluency is a growing and developing competency using a system realized through sound, what instances do we see of it in our socio-ec(h)o study?

5.2. Dimensions of Aural Fluency

While the statistical results above attempt to address, albeit in a limited way, the measure of effectiveness of the auditory display as it informs, helps and directs participants within an ambient intelligent setting, theorizing about the auditory display points to aspects of conceptual knowledge about the role of sound feedback in such environments that we aim to explore and help define. Building upon the emergent notion of *aural fluency* presented, we articulate a set of preliminary heuristics about auditory display that stem from our socio-ec(h)o study.

We suggest that *aural fluency* is a level of competency achieved with and supported by dynamic, ambient and continuous sonification. Like a language, the audio is required to provide a consistent and intelligible structure and expression. Unlike a language, the issues are less complex but also the result of conscious design. There is the need to improve design, however the resulting audio must be credible in order to engage listeners to learn it. We offer four heuristics that help in the design of *aural fluency:*

Epistemic Dimensions: The way players had to make sense of the audio feedback was through both an analytical and embodied involvement. While balance in this respect is critical, a degree of ambiguity plays a role in motivating listeners to learn and understand. Many of the groups reported being on many occasions confused or unsure about the sound feedback, however it was often a matter of degrees of 'clarity' rather than complete confusion; and teams did not seem to want to quit because of it, in fact it seemed to invite more attentive listening and feeling of challenge. This concept relates to the notion of everyday listening that people are already proficient in from interacting with their natural environment – making sense of confusing, unclear, complex sonic situations by selectively focusing or shifting attention on different aural elements, and fine-tuning their ear to particular sound structures and qualities. This fine-tuning, in turn, relates to our idea of aural fluency.

Narratological Dimensions: Leveraging past associations and creating new ones mediates expectations of what particular sounds mean and how to engage with them. As mentioned, the concept of aural fluency helps make sense of findings that show a positive experience of the system and appreciation of the informative qualities of the auditory display in the face of long and inconsistent completion times. Further, when examined through the lens of everyday listening practices, as developed by Schafer and Truax [22], we see numerous examples of both players forming a narratological association with the soundscape – "no, no we got no fire, somebody has to keep the flame!"… as well as a seamless connections to embodiment – "so the fire builds up and we're all still…what if we move towards the sound? What if… when the fire gets all crazy we have to move more with it?" This narratological relationship could also serve to explain why negative polarity was not a problem in Level 2, where the solutions required moving very slow (so the slower they move the slower the sound gets) but it was a problem in Level 5 where the relationship between puzzle, movement required and sound feedback was more abstract.

Communicational Dimensions: In AmI systems the environment – in our case primarily comprised of sound - represents the means of accessing and understanding the system's intelligence, rules and behaviour. This concept arose out of an observation we made that players would often anthropomorphize the system via its sonic response – 'no, not like this, it doesn't like that, it gets all quiet'. Sound feedback appeared to be something that users saw as the physical manifestation of the system and relied on it heavily in order to understand how it works and responds to them. In a way it showed a conversational style of experimentation and eliciting response, in order to learn to manipulate, interpret and be in the environment. Again, this concept speaks to the idea of everyday listening [2, 22] – selectively paying attention to minute sound shifts to gain information about the soundmaker's state, thus building up aural fluency. It brings a different perspective to auditory display design to think of sound feedback as an epistemological occurrence in itself - the content and form in which participants understand and think of the system, which is otherwise ambient, invisible and yet physically persistent. This notion in turn connects the perception of competency that players build up from manipulating the system with their aural fluency, also developed over time. Once again, this helps to account for why even teams who took a very long time to complete a level felt competent in mastering the system's response and positive towards the sonic feedback.

Confirmatory Dimensions – We observed that having feedback at all times, in this case continuous soundscape feedback that responds with subtle yet perceivable changes to group actions, reinforces and rewards the efforts of interacting with the system. It provides clear scaffolding for achieving *aural fluency.* Our case of intensity-based gradient soundscape most resembles a sonification scenario where the response is constant and dynamic, and requires constant attention. The anticipatory reward sound – signaling that users are close to completion served as a soundmark, which users became quickly familiar with, and could instantly identify so as to adjust their actions accordingly. The closer they were to a solution the more time they spent in acute listening attention to the anticipatory reward sound. Upon completion of each game level the auditory reward sound – a signal in the soundscape - seemed to be really enjoyed by the players, and important to their leaving the system confident and satisfied, rather than frustrated and defeated.

6. CONCLUSIONS

Despite the vast proliferation of audio-augmented technologies and spaces, relatively little work has been done to date in designing multi-layered informational auditory displays for responsive environments that actively guide human activity towards solving a problem or achieving a goal [24, 16] While intensity-based sound feedback occurs normally in particular everyday circumstances (e.g. paging your cordless phone at home and going from room to room listening to its sound intensifying), there are few studies of these everyday phenomena, and fewer still of their possible translation into design guidelines for sound feedback. Furthermore, few studies focus on perception of complex everyday changing sound, while taking into account context and purpose of activity, level of embodiment, or familiarity and associations with the sound [24, 25, 26]. A methodological investigation into all these different components of the model is needed in order to understand more fully how to better use and design such auditory displays. Finally, the demands and unique requirements of ambient intelligent systems are still in the process of being heuristically and methodologically developed by the human-computer interaction and adjacent communities, which makes for challenges at this point in time to uncovering stable auditory perception and auditory display design guidelines. What we have presented here is an example of a design-based research project attempting to merge design methodologies with auditory display design investigations. Our analysis starts from the empirical data and opens itself up to emerging research questions and alternative explanations for our results. From that standpoint, we develop and discuss the notion of *aural fluency*, as well as build up and describe several general heuristics related to the role of ambient, informative auditory display in ubiquitous computing settings.

7. REFERENCES

[1] Adcock, M., Barrass, S. (2004)."Cultivating Design Patterns for Auditory Displays." In Proc. ICAD 04.

[2] Gaver, W. (1994). Using and creating auditory icons. In G. Kramer (Ed.), Auditory display: Sonification, audification, and auditory interfaces (pp. 417-446). Reading, MA: Addison-Wesley.

[3] Walker, B., & Kramer, G. (1996). Mappings and metaphors in auditory displays: An experimental assessment. Presented at the third International Conference on Auditory Display (ICAD 1996), Palo Alto, California.

[4] Brewster, S.A., Wright, P.C. & Edwards, A.D.N. (1992). A detailed investigation into the effectiveness of earcons. In G. Kramer (Ed.), Auditory display, sonification, audification and auditory interfaces. Santa Fe, NM: Addison-Wesley, pp. 471-498.

[5] McAdams, S. & Bigand, E. eds. Thinking in Sound: The cognitive psychology of human audition, Clarendon, 1993.

[6] Lepletre, G., & McGregor, I. (2004). How to tackle auditory interface aesthetics? Discussion and case study. *Proceedings of the Tenth Meeting of the International Conference on Auditory Display (ICAD),* Sydney, Australia.

[7] Butz, A., & Jung, R. (2005). Seamless user notification in ambient soundscapes. *Proceedings of the International Conference on Intelligent User Interfaces (IUI 2005)* (pp. 320-322). New York: ACM Press.

[8] Hunt, A., & Hermann, T. (2004). The importance of interaction in sonification. *Proceedings of the Meeting of the International Conference on Auditory Display (ICAD),* Sydney, Australia.

[9] Gibson, J. J. (1966). The Senses Considered as Perceptual Systems. Boston: Houghton Mifflin.

[10] Nardi, B. (Ed.). (1996). *Context and consciousness: Activity theory and human-computer interaction.* Cambridge, MA: MIT Press.

[11] Svanæs, D., & Seland, G. (2004). Putting the users center stage: Role playing and low-fi prototyping enable end users to design mobile systems. Proceedings of the SIGCHI conference on Human factors in Computing Systems (CHI '04) (pp. 479-456). New York: ACM Press.

[12] Sanders, E.B.-N. (2001). Collective creativity. Loop: AIGA Journal of Interaction Design Education, 3(7). Retrieved May 15, 2007, from http://loop1.aiga.org/content.cfm?Alias=sandersucd

[13] Wakkary, R., Hatala, M. Lovell, R, Droumeva, M., Antle, A., Evernden, D. & Bizzocchi, J. (2005). socio-echo: Ambient intelligence and gameplay. In *2005 International conference: Changing views: Worlds in play*. Burnaby, BC: Digital Games Research Association.

[14] Droumeva, D. & Wakkary, R. (2006). The role of participatory workshops in investigating narrative and sound ecologies in the design of an ambient intelligence audio display. In *Proceedings of the 12th meeting of the international conference on auditory display* (pp.36-43). London: ICAD.

[15] Walker, B., J. Lindsay and J. Godfrey. (2004). The Audio Abacus: Representing Numerical Values with Nonspeech Sound for the Visually Impaired. Presented at ASSETS '04 in Atlanta, Georgia, ACM Press.

[16] Droumeva, M., & Wakkary, R. (2006). Sound intensity gradients in an ambient intelligence audio display. In *CHI 2006: Conference on human factors in computing systems* (pp.724-729). New York: ACM Press.

[17] Flowers, J. H., Whitwer, L. E., Grafel, D., & Kotan, C. A. (2001). Sonification of daily weather records: Issues of perception, attention, and memory in design choices. Proceedings of the International Conference on Auditory Display (ICAD 2001).

[18] Hermann, T., Meinicke, P., & Ritter, H. (2000). Principle curve sonification. Proceedings of the International Conference on Auditory Display (ICAD) (pp. 81-86).

[19] Nesbitt, K. V., & Barrass, S. (2002). Evaluation of a multimodal sonification and visualization of depth of market stock data. Presented at International Conference on Auditory Display (ICAD), Kyoto, Japan.

[20] Smith, D. R., & Walker, B. N. (2002). Tick-marks, axes, and labels: The effects of adding context to auditory graphs. Presented at International Conference on Auditory Display (ICAD 2002), Kyoto, Japan.

[21] Cook, P. (Ed.). (1999). Music, cognition and computerized sound: An introduction to psychoacoustics. Cambridge, MA: MIT Press.

[22] Truax, B. (2001). Acoustic communication (2nd ed.). Westport, CT: Ablex.

[23] Löwrgen, J. (2007) Fluency as an Experiential Quality in Augmented Spaces, International Journal of design, Vol. 2 (2007) pp. 1-10.

[24] Takahata, M., Sakane, Y., & Takebayashi, Y. (2004). Sound feedback for powerful karate training..In NIME '04: Proceedings of the 2004 Conference on New Interfaces for Musical Expression (pp. 13–18), Singapore.

[25] Parente, P., & Bishop, G. (2004). BATS: The Blind Audio Tactile Mapping System. In Proceedings of ACM South Eastern Conference (ACMSE), Savannah, GA.

[26] McElligott, J. (2004). Designing sound tools and toys for blind and visually impaired children. *Proceedings of the 2004 Conference of Interaction Design and Children (IDC)* (pp. 65-72). New York: ACM Press.