

Creative Evaluation

Georgios Marentakis

georgios.marentakis@tugraz.at
Signal Processing and Speech Communication
Laboratory, Graz University of Technology
Graz, Austria

David Pirrò and Marian Weger

{pirro,weger}@iem.at
Institute of Electronic Music and Acoustics
University of Music and Performing Arts Graz
Graz, Austria

ABSTRACT

Interactive systems are traditionally evaluated against an ‘intended’ use by involving external participants. This approach has been challenged recently because of difficulties in addressing applications without an ‘intended’ use or an ‘intended’ interpretation, but also because the propositional-ity of the evaluation medium may not address the aesthetics of interactive systems sufficiently. We turn our attention to the evaluation of interactive art, in which, although both difficulties emerge, traditional evaluation methods are commonly used. In trying to stay open to interpretation and address aesthetic thinking and knowledge, we introduce and apply creative evaluation. Ten artists were asked to both direct themselves enacting their interaction experience and to express it using artistic media. Inspiration was obtained in two interactive installations. The results of this experiment demonstrate the ability of artistic practice to maintain interpretation variability and its capacity to address aesthetic thinking and knowledge in the evaluation of interactive systems.

ACM Classification Keywords

J.5 ARTS AND HUMANITIES; H.5.2 User Interfaces: User-centered design; H.5.2 User Interfaces: Evaluation/Methodology; H.5.5 Sound and Music Computing: Methodologies and Techniques; H.5.1 Multimedia Information Systems: Evaluation/methodology

Author Keywords

Authors’ choice; of terms; separated; by semicolons; include commas, within terms only; required. User Evaluation, Interactive Art, Artistic Research

INTRODUCTION

Historically, evaluation in HCI has been synonymous to *usability engineering*. Usability engineering models interaction as the communication between the designer and the user through a proxy i.e. the interface. This allows usability to be quantified using information-theoretic measures, such as the

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

DIS 2017 June 10-14, 2017, Edinburgh, United Kingdom

© 2017 Copyright held by the owner/author(s). Publication rights licensed to ACM. ISBN 978-1-4503-4922-2/17/06... 15.00

DOI: <http://dx.doi.org/10.1145/3064663.3064710>

interaction bandwidth, error rate, and error recovery. These measures may be applied to characterize communication at different levels of abstraction within an interface that cover the whole range from the mental model of its operation to its ergonomics. Perception, cognition, and action are treated as separate interacting blocks, as in the highly influential Action Cycle [32] or the Keystroke-level model [13]. External participants are necessary in the evaluation in order to obtain the measurements required to characterize and optimize usability.

This ‘information-theoretic’ view is evident in the ISO definition of usability in which usability is defined as the efficiency, effectiveness, and satisfaction with which users can achieve their goals by using a system in a specified context of use. Effectiveness and efficiency have been assessed primarily by measuring task-completion-rate and task-completion-time in closed-task laboratory experiments [27]. User satisfaction has been assessed by mixed techniques such as Likert scales, user observation, and the analysis of videos and interviews [27].

Eventually, it has been acknowledged that user satisfaction (and its hedonic quality) is only one of the many dimensions of user experience¹ (UX). Other dimensions are emotions and affect, enjoyment, and aesthetics [1, 34]. Furthermore, the embodied, situated, and social aspect of experience needs to be considered [15]. The multidimensional and context-dependent nature of UX makes it difficult to quantify. For this reason, methodologies which originate in qualitative research techniques [35] and take place in the field using non-task-oriented scenarios have been used in order to assess UX [1]. The turn to UX research essentially acknowledges that, although necessary, usability may not be sufficient for a good user experience.

The shift in evaluation methodology required in order to investigate user experience is especially important for the design and artistic research practices in interaction design. This is because both design and artistic works aim at eliciting a variety of interpretations which need not converge in a statistical sense or conform to a specification [23, 22, 37, 21, 20]. While therefore usability and UX methods are still relevant [12, 11], their application is complicated because the traditional evaluation goal of quantifying and bridging

¹The ISO definition of user experience is “a person’s perceptions and responses that result from the use or anticipated use of a product, system, or service”.

the gap between designer and user interpretations is only partially relevant.

A further difficulty originates in the importance of aesthetics for the design and artistic research practice. Addressing aesthetics using user experience and usability research methods is difficult, because such an investigation needs access to aspects of the experience that may not be easy or even possible to verbalize or quantify [7].

Applying evaluation within the design and artistic research practices, requires therefore devising novel evaluation methods that will allow evaluators to proceed beyond contemporary usability and UX research methods.

In this paper, we attempt to design such methods in the context of the evaluation of two interactive sound installations. We do not aim to assess the value of these artistic works. Rather, we research evaluation techniques that can address interpretation variability and aesthetics and in this way contribute to their user-centred design. Although we focus on interactive art, the outcome of our investigation may also be relevant for the evaluation of interactive systems in general.

EVALUATION OF INTERACTIVE ART

Interactive artworks have been evaluated primarily by applying usability and UX research methods. These methods have allowed artists to engineer their usability [26, 12, 17] and to obtain insight into visitor experience [12, 17, 5, 4]. Numerical, video, and verbal data from evaluations have been used for modelling visitor behaviour [10, 6], in documentation activities [24], and for providing feedback to curators and funding bodies² [16, 29, 6, 31].

How to extend evaluation beyond the practices mentioned above? After all, artworks are meant to operate at deeper levels, and endow insights into emotions, human nature, relationships, and our place in the world [36]. Furthermore, artworks deal with aesthetics. How can evaluation address these aspects and provide constructive feedback beyond simple judgments of appeal?

The first difficulty relates to the fact that although artists are usually clear in their practice, it is rather rare that they will formulate an explicit intention with respect to the aesthetics or the impact of their work. This should not be taken as a sign of unclear aesthetics or motivations. Simply, artworks are not, as a rule, conceived in order to support a single interpretation [21], nor do they try to communicate a certainty, or a fact [36].

Accordingly, it is also rare that art may be experienced in a uniform way. As if by design, the resulting experience is one of potentialities. Visitor interpretation is therefore more likely a hypothesis out of the multiple potential meanings, rather than a conclusion [36, 8].

The traditional approach of evaluating against something is hindered therefore both by the lack of such concrete views

²We do not intend to imply here that these may suffice to evaluate the artistic value of a work.

on the side of the artists and by the diverging ways with which the artworks may be received by the audience.

Aesthetics in interaction design have been seen as a way to promote bodily experiences, complex symbolic representations, and to create involvement, experience, surprise, and serendipity in interaction [33]. This experiential view of aesthetics draws on pragmatist aesthetics (cf. Dewey [14] and Shusterman [38]) to stress the situated nature of art as experience.

Based on these ideas, Wright et al. [39] created their account of aesthetic experience. Their intention was to understand people's interactions and relations with technology, and to engineer user experience. Their account is holistic and encompasses all of the sensory, emotional, spatio-temporal, and compositional threads that constitute user experience. It does emphasize continuous engagement and sense making during interaction, and acknowledges anticipating, connecting, interpreting, reflecting, recounting, and appropriating as integral parts of sense-making. It promotes a dialogical view, in which experience is understood in relation to the state and the context in which individuals find themselves while engaging with technology.

Although this approach to 'designing' aesthetics may be entirely relevant for the design of products and experiences, it appropriates aesthetics to achieve specific goals. Artists, however, do not aim at eliciting emotional, cognitive, or other responses [3]. In art and artistic research, aesthetics are both the object of the artists' research and the tool with which artists perform their research [30].

Unfortunately, it is inherently difficult to elicit a discourse through evaluation at this level because of the propositional nature of the statements that we have learned to expect as the outcome of an evaluation. Although propositional statements may comment on the *aesthetic experience*, they do not entail it nor can they provide us with the subtlety of the moment in which *a person is moved by an artistic work*. It is therefore important to investigate how evaluation should be designed to address aesthetic thinking and knowledge. But what is aesthetic thinking and aesthetic knowledge in the first place?

AESTHETIC THINKING AND AESTHETIC KNOWLEDGE

The nature of aesthetic thinking and aesthetic knowledge has been the object of intense philosophical inquiry. Their definition would be the subject of an epistemology of aesthetics, whose formulation has occupied philosophy since the antiquity. Such an epistemology is important to us, because it could serve as the foundation for the design of techniques for the evaluation of aesthetic thinking and knowledge.

The topic is receiving revived interest as art is integrated deeper in academia and artistic doctorates become common [9]. Pretty much similar to the discussion on design research, the argument goes that research should generate knowledge, therefore artistic research should generate artistic knowledge. Knowledge requires thinking – in our case artistic thinking. Does artistic thinking and knowledge relate to scientific or philosophical thinking and knowledge? Fundamental

differences between art, science, and philosophy may lead us to believe that this is only rarely the case.

Science and philosophy aim to create generalized theories. Art, instead, is the science of the singular. Artistic positions formulate their own model, none of which is comparable to each other. [30, 8, 9, 36].

Scientific and philosophical statements strive to become unambiguous. Instead, the experience of art is one of *potentialities*. This makes visitors' interpretation more of a hypothesis than a conclusion [36].

Scientific and philosophical statements need to be justified. By contrast, artworks are hardly ever concerned with providing justification. Whatever knowledge is generated upon exposure to art, is therefore the outcome of an inductive rather than deductive process [40].

Taken together, these arguments show that artistic thinking and knowledge do not share much with their scientific or philosophical counterparts. The broad consensus is³ that art is not, as a rule, a source of knowledge in the way knowledge is understood in science or philosophy [30, 2, 8]. This does not mean that art cannot be the source of knowledge, it rather points towards the existence of a non-propositional element in artistic thinking and knowledge.

Young identifies this to be *practical knowledge* [40]. Practical knowledge originates in the potential of art to teach us how to do something or provide us with the ability to recognize something and can be accumulated in the absence of explicit (theoretical) knowledge. The transition from practice to theory is not always possible, sometimes due to the limitations of language. Biggs [2] defines it as knowing-how, in contrast to knowing-that.

Borgdorff [8] identifies it as the “unreflective, non-conceptual content enclosed in aesthetic experiences, enacted in creative practices, and embodied in artistic products and practices”. The cognitive value of art is therefore its ability to allow us to “sense something of our pre-reflective intimacy with the world, while realizing that we will never explicitly understand what lies there in such plain view”. This is a mechanism which sets thinking in motion, “an invitation to unfinished thinking”. Thinking in artistic research is the articulation of these “contingent perspectives”, which can be understood as the articulation of embodied knowledge [9].

Mersch [30] links the production of knowledge in artistic research to the production of aesthetic knowledge, which presupposes a “thinking in the arts” or “thinking in aesthetics”. Exposure to art elicits something, which otherwise cannot come into being and leads to thinking and reflecting.

³A mild exception is Young who argues both ways [40]. Art could be the source of propositional (but also non-propositional knowledge) and improve both our understanding and our judgment. The communication of both types of knowledge requires the convincing demonstration of the *rightness of a perspective*. Perspectives are communicated using the representational power of artistic illustration. They need not be justified. Their *rightness* is demonstrated in the artwork; their truth, however, is the subject of empirical inquiry.

Importantly, although an engagement with aesthetics appears in every design practice (e.g., in decoration), the notion of art and its epistemological impulse is associated with its ability to set thinking in motion. Mersch [30] identifies artistic thinking as a praxis or even as a performance.

Both Mersch and Borgdorff agree that artistic thinking and the generation of knowledge in research through the arts [18] occur without the need for reasoning or causality and are non-predicative and pre-reflective, as we would assume for intuition or imagination. Furthermore, according to Mersch, not only is artistic thinking neither philosophical nor scientific, it *cannot* be translated in a propositional medium.

A particularly relevant contribution of Mersch [30] is in identifying *showing*, rather than saying, to be the main epistemical mode of art. The term self-showing (sich-zeigen) is used to refer to the use of the artistic medium to investigate itself. The notion of art as deixis can be used to explain the multiplicity of ways in which art can be perceived [2, 40].

CREATIVE EVALUATION

It is likely clear by now why the propositional nature of the usability and user experience evaluation methods (e.g., the audiovisual-cued recall⁴ [12], verbal protocols in general, scales, or video recordings) that have been used for the evaluation of interactive art, can only partially contribute to the evaluation of aesthetic thinking and knowledge.

In dealing with the ineffable⁵, Boehner et al. [7] opted for multimodality in the evaluation materials. They confronted themselves on a regular basis with shared journals, diverse materials, and thoughts, and freely switched between roles; designer, developer, and evaluator. This enabled them to reflect on their research object⁶ in different ways and come up with important realizations. Despite the obvious value of such dynamic feedback methods [37], they only partially address how to maintain a discourse at a pre-reflective level.

Creative evaluation advocates the use of creative practices in the context of interactive systems evaluation. Even more, creative evaluation suggests that participants articulate their interactive experience using artistic practice. Involving artists in the evaluation is important not only because they are experts in accessing non-propositional layers of experience, but also because they are trained in eliciting aesthetic thinking and knowledge.

The concept emerged as a reaction to the difficulty of evaluating against an ‘intended’ use in interactive artworks and the substantial discussion on the non-propositional nature of artistic thinking and artistic knowledge in the literature. It is essentially a turn away from the metrics-driven evaluation, towards maintaining freedom of interpretation while addressing aesthetic thinking and knowledge. The aim is to enrich existing evaluation techniques.

⁴a process in which participants described what they were experiencing in the installation based on information in the videos

⁵in this case emotion

⁶a system for the display of affective information

Employing art for qualitative research purposes is not new. Especially within the social sciences, there has been a notable effort to incorporate art as a research method [28]. However, we are not aware of any applications of artistic methods, neither in interactive art evaluation nor in interactive systems evaluation. Interestingly, applying art to evaluate, and essentially research art, would constitute artistic research [30]. Involving artists in interactive experience evaluation is different from the cultural commentators concept [19]. There creative practitioners are invited in order to help with the interpretation of data gathered from using cultural probes, or to document the experiences of people living with prototype designs in their homes. Therefore, they do not evaluate the works themselves.

EXPERIMENT

We based our experiment with the design of creative evaluation on ‘showing’. Showing (in contrast to saying) has been identified as the main epistemical mode of art [30].

Specifically, we asked participants to *show* what they considered important about the artworks that we evaluated in two different ways. The first required participants to compose, direct, and perform a short video in which they enacted their experience of the installation by interacting with it. The second required participants to use their preferred artistic medium (i.e., the one in which they had been trained) in order to create an artwork inspired by their experience of the installation.

Procedure

The evaluation was executed in three different sessions. In the first, participants explored the installations alone and unobserved. In the second, they directed a video in which they showed *what they believed was worth communicating about their experience with the installation*. In the third, they performed a work (in four out of five cases explicitly an artwork). This was conceived by themselves as a response to their experience with the interactive art installation. Both the video and the artwork were documented. All works were presented together with the installations during a *finissage* public event. Subsequently, participants were asked to write two paragraphs as a commentary to their video and artworks. They had a week’s time to return these short texts.

We evaluated two interactive sound installations, each by a different participant group. Sessions were distributed over one week and were recorded. Evaluation took place in the *esc medien kunst labor* in Graz⁷. Each session lasted 1.5 hours. Participants stayed for as long as they liked and were free to ask for as much information as they considered necessary. Participants were briefed on the whole process during the first session. They were allowed to take notes during the different appointments in anticipation of the next evaluation steps.

The installations were staged for public viewing. With the exception of the *finissage*, the evaluation sessions took place outside the public exposition hours.

⁷<http://esc.mur.at/>

Installations

The interactive sound installations ‘Random Access Lattice’ by Gerhard Eckel and ‘Paris Flâneur’ by Martin Rumori were evaluated. A graphic illustration of each installation is provided in the top row of Figure 1.

Random Access Lattice uses a loudspeaker on which reflective markers are attached. This allows the real-time tracking of its location by a state-of-the-art optical tracking system. The loudspeaker is silent when still and plays back sound when moved within a specified area, this of a surrounding cube. The moving loudspeaker can be interpreted as the head of a tape player. It plays back sound material that is spatially arranged in straight lines on a three-dimensional grid.

The sound material consists of poems recited in 40 different languages, together with their copyright notice. Grid lines are parallel to one of axes of the surrounding cube. A rectangle on the floor marks the bottom side of the cube, which extends a bit higher than most people would be able to reach.

The grid is rather tight and has a large number of subdivisions. As a result, the same sound material may be found in different areas of the grid. Poem recitations are clearly audible only when one moves slowly directly along a grid line. Otherwise, the sound of neighbouring lines is used to create a sonic outcome that varies depending on loudspeaker speed and exact location. Moving back and forth creates loops, whose sonic variability is increased when deviating from grid lines. Faster movements result in louder sound.

Paris Flâneur is an interactive auditory virtual environment that uses binaural rendering and tracked headphones. Rendering is done using generalized HRTF functions. Headphones are equipped with markers whose location and orientation is tracked a state-of-the-art optical tracking system.

In total, seven invisible but audible ‘sound islands’ are arranged in the virtual auditory space. Listeners can find them by approaching, turning, and listening. The ‘islands’ represent sound situations of everyday life. They have been recorded binaurally in and around the city of Paris.

If the listener stays at the same position for a while, the sound islands will begin to move towards him or her. When located very close to the centre of an island, the listener will sonically ‘enter’ it and gain a different audible perspective, while the other islands will become silent. Moving along, the surrounding island will follow for some time; once listeners increase their speed, it will be left behind and the virtual scenery will fade in again.

Participants

Five participants evaluated each installation. They were either post-graduate students⁸ or professionals in the fields of design, music, dance, music aesthetics, or computer music. Five were female and five male.

⁸of the University of Music and Performing Arts Graz or of the University of Applied Sciences FH Joanneum

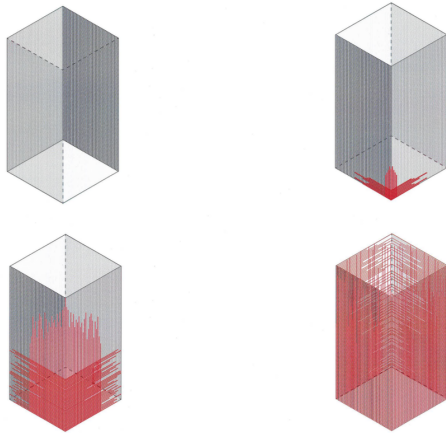
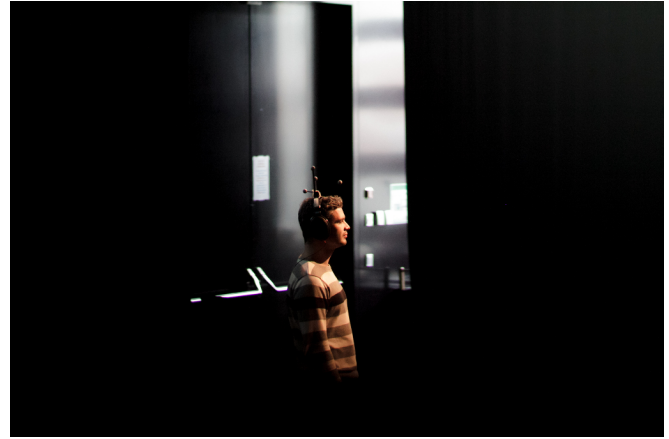


Figure 1. Top: Two views of the sound installations, Bottom: the designer's works as a result of experiencing the installations. Left column: Random Access Lattice, Right column: Paris Flâneur.

Within the second session, all participants finished the 'showing' video using only the installation. Within the third session, each designer created a graphic work, each musician (a violin and a double-bass professional) a musical performance, each dancer a dance performance, each music aesthetics professional an essay, and each computer musician a computer music work. Participants worked alone and were advised that works and videos should not exceed five minutes in duration. Participants received a monetary compensation for their time and efforts.

Results

All participants created videos and presented an artistic work. Unfortunately, certain pre-reflective and aesthetic aspects of the created materials may only be accessible by directly experiencing them. What we offer next is a textual description of their content.

Random Access Lattice

In her video, the designer focused on the linearity of the audible space created by the installation and on the skill required in following grid lines to create comprehensible sound. She exclusively performed movements on the outer surface of the (virtual) cube that enclosed the installation, but not inside. Her movements were parallel to the three

Cartesian axes. She started by using cautious slow movements and gradually increased their speed. She used the rectangular area marked on the floor to guide her movement. Her design work was a 3D visualization of a cube overlaid by grid-forming lines of different thickness. In her text, it is clear that she was highly concerned with the spatial structure of the installation. She considered the alignment of the speech tracks along the three-dimensional grid lines as the most important thing to communicate.

In his video, the computer music artist demonstrated the sounds that could be created by minuscule loudspeaker movements around a given position or by inducing deviations in the calculated loudspeaker location by hiding visual markers. He contrasted them to the sound of the recited poem excerpts that emerged due to the slow careful movement of the loudspeaker along the grid lines. To create his piece and performance, the computer music artist developed a variation of the installation. First, he replaced the speech tracks in the installation with percussive sounds. Then, he modified the way loudspeaker movement created sound. The result was that it was easier to create reproducible sounds using the installation. In his performance, he largely replicated the movements he used to create his video; an evidence to

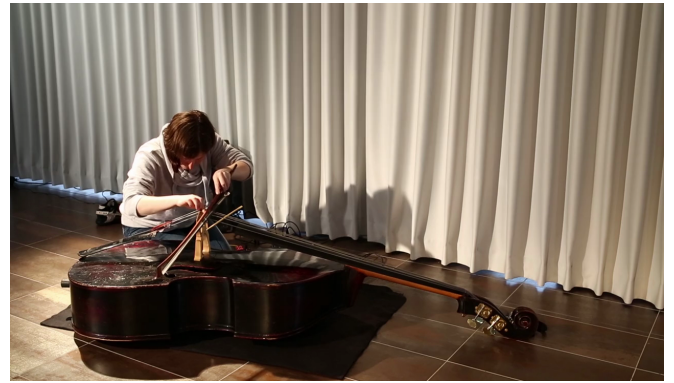


Figure 2. The participants during the performance of their works. From top to bottom: musicians, dancers, music aesthetic professionals, computer musicians. Left column: Random Access Lattice, Right column: Paris Flâneur.

the difficulty of creating reproducible sounds in the installation.

In his video, the violin player also used the sounds that could be created by minuscule loudspeaker movements around a given position or by inducing deviations in the calculated loudspeaker position by hiding the markers used by the tracking system. He also contrasted them to the sound of the recited poems and to the sounds that could be created by fast movements of large amplitudes. His hand and body movements were very diverse. He wrote that he found the sonic possibilities that emerged when playing the installation like an instrument to be the most important thing to communicate and that he used his time to develop a playing technique that he could reproduce in his video. To create his piece, the violin player mapped the body of his violin to the space spanned by the installation and the bow to the loudspeaker. He then tried to demonstrate sound variations by moving the bow at different locations. He mapped the tension he experienced while playing the installation to the tension he used at his bow. He mentioned that he would have appreciated having had more time to refine the aesthetics of his work.

The dancer structured his video around retrieving a specific sentence that could be found in several specific spatial positions. Despite having kept notes on their location, this was not entirely easy. He commented that this difficulty could be traded against the novelty of what may pop up while searching. His dance piece was a dialogue between body and hand postures. It studied the ways with which these could develop while interacting with the installation. The constraints posed by the linearity of the grid system underlying the installation were evident in his movements. He explicitly mentions having played with the idea of a “graphic body, keeping in mind the way the grid is constructed and how his body parts and angles relate to it” and also “having played with the idea of becoming the loudspeaker and then gradually morphing into the surrounding space”.

In his video, the music aesthetics professional started with careful hand and body movements which he performed while listening carefully to their result. Subsequently, he revisited specific areas and the sounds that resided there. He proceeded by introducing repetitive movements that created novel variations of the sonic material in a given location. Later in the video, his movement speed and confidence increased. In his text, he clearly mentions that he tried to demonstrate his learning process in three steps: a. listening and moving, i.e., the exploration process, b. understanding, i.e., making sense of the sound material in the grid, and c. creating, i.e., using the installation in a creative way by creating variations through repeatable movement patterns. He tried to describe these three phases with words in his essay. He composed the text in a free style which resembles poetry rather than prose.

Paris Flâneur

In her video, the designer moved through sound islands and demonstrated arriving in them vs. departing back to the complex soundscape. She often closed her eyes, lingered

in places, walked with the sounds, moved towards or away from sounds, ignored some, or moved rhythmically to others. In her text, she brings forward: a. the immersion she experienced with the binaural system, b. the evocation of a feeling of Paris, and c. the structured sound-walk strategy she used in her video. The latter aimed to construct a narrative of specific thoughts and emotions. Her artistic work was a collage which featured images she associated with the sounds in the installation. She said that she chose collage because it offers the possibility to arrange the associations brought forward by the sound material used in the installation spatially.

The computer music artist started his video with a ‘sound-walk’ through the installation. He then stood still and waited for the sounds in the surrounding binaural recordings to approach. Eventually, he moved into one recording and walked away slowly while listening to it. He mentioned that despite small difficulties with identifying the nearest sound, which occasionally led to unwanted sounds approaching him, he was able to execute his planned actions. To create his performance he constructed a Wii controlled interface that allowed the browsing and play-back of the recordings comprising the installation. In an audio game fashion, he then strolled towards sounds of interest in the installation and synced his Wii-controlled tracks to them. Both were mixed together in his headphones. During the performance, he installed two speakers that would present the audience with noise until he had achieved this task, which he repeated a number of times. He stated that he was inspired by the contrast between the sonic variability of the installation to the silence associated with externally observing people interacting with the installation.

In her show, the double-bass player moved fast and created an audible contrast between the binaural recordings occupying different spatial locations. She then picked a recording and moved slowly with her eyes closed while listening. She used her hands to sense and avoid obstacles and then sat on her knees and remained still while listening. At last, she tried to find her way back to the first recording in her walk with her eyes closed. She mentioned that headphones and cables often stood in her way. She claimed that when her eyes were closed, the spatial dimension of the sound material altered her sense of orienting. To create her piece and performance, she used a contact microphone to obtain the sounds created by exciting different parts of her instrument (bridge, strings, ...). This created an association between spatial location and sound. She placed her instrument on the floor to ease her task. Using live-looping she then post-processed and overlapped these sounds electronically. This allowed her to even sketch her favourite sound in the installation.

In her video, the dancer also started with ‘soundwalk’, augmented with movements in response to the content of the sound recordings at the different locations. Sometimes the link between sound and movement was explicit, i.e., looking up to the sound of a helicopter going by, while occasionally, the association was free. In her text, she added that she spent more time in aurally pleasing areas and walked away from less interesting ones. Her dance performance was inspired

by the spatiality of the installation, the associations evoked by the sounds, the ‘hopping’ between sound islands, and the possibility to move sounds around. She reinterpreted these aspects a movement repertoire which she used to compose her piece. She noted that her dance piece offered insight in her experience and in no means was it a demonstration of how the installation works.

In her video, the music aesthetics professional also started with a ‘sound walk’, noticing some sounds in the recordings, returning to them, or walking away from others. She then ‘entered’ a recording and stood still. Consequently, she moved slowly away while listening to the sound. Then she stepped back to the soundscape and repeated the whole process using other recordings in the installation. The result was that she rearranged recordings to different areas. In the end of the video, she stood next to an area that matched her intentions. She described her intention to be the spatial rearrangement of the binaural recordings in order to create an aurally pleasing area. She mentioned that this was not entirely possible to achieve and provided an explanation. Her essay was inspired by the impossibility of finding silent spots in the installation⁹. This brought up associations of an “urban mess”, a feeling of being “alone but never unaccompanied”. Her essay very much stresses the experience of moving and listening, the relationship between sound and space, and the near-impossibility of silence. The essay was not structurally formal.

DISCUSSION

In our effort to stay open to interpretation and address the pre-reflective level of experience in which aesthetic thinking and knowledge operate, we devised and applied creative evaluation. The outcome of the evaluation was the videos in which participants enacted their experience with the installation and the artistic works they have created. These allowed us, as well as the *finissage* audience, to experience at a pre-reflective level what the participants considered worth communicating.

This non-verbal experience not only stimulated us mentally and emotionally, but also aesthetically. We believe this multiplicity is especially interesting because it allows an observer to see a given work from perspectives that are otherwise not available. This would have been the case if categorical statements monopolized the evaluation outcome, for example.

The ways with which our exposure to the works and the texts the participants provided us with, stirred our thinking and imagination are summarized in this section. A commentary on the epistemological status of this feedback is outside the scope of this paper. We hope, however, that its relevance for the installations is easy to spot.

Categorical feedback about the installations

Concerning Random Access Lattice, the following aspects were emphasized in the evaluation outcome: a. the contrast

⁹even in quieter spots sounds would eventually start moving towards the visitor once standing still

between the sound of the recited poems as the result of orderly movement versus the richness of the sonic outcome associated with other movements, b. the linearity of the space in which the installation lived, and c. the playability of the installation as an instrument.

A major concern was the difficulty of creating reproducible sounds. This difficulty can be attributed to the interaction technique used and to the tightness of the grid subdivisions. Creating the same sound would require repeating the exact loudspeaker movement in the exact same position of the 3D space. The difficulty inherent in repeating an exact movement trajectory with a physical object in 3D space, together with the minimal spatial reference points in the installation made this a difficult, but not impossible, task.

Upon discussing with its creator, it was made clear to us that this was a conscious design choice whose intention was that the installation resists a single interpretation. He stated: “evidently, evaluation depends on the value system which is in force in a field. The traditional values of usability (immediate understanding and gratification, no learning needed, etc.) do not necessarily make sense for artworks. An artwork may in itself be critical of these values. This is the case with Random Access Lattice, which defines a particular resistance, the structure of which (which is what I did compose) defines the identity of the piece and strongly shapes its experience (or the type of experience I am interested in).”

Concerning Paris Flâneur, the act and experience of a ‘sound-walk’ appeared predominantly in the evaluation outcome. Furthermore, the interaction techniques in the installation were often appropriated for the construction of spatial sonic narratives. The content of the binaural recordings touched participants at different levels of experience. It stirred emotional, gestural, and conceptual associations. These permeated the artistic works which were ‘rich’ in this respect.

In their works, participants very often contrasted the experience of listening to a single binaural recording (while on an island for example) to this of moving around the space occupied by all recordings. The artistic exploration of this contrast has been explicitly formulated as a research topic by the artist.

The ability to pull sounds by standing still did also receive attention in the evaluation outcome. Together with re-placing sound islands, it was used to assist in the creation of spatial sonic narratives and was appreciated by the users. Spatial sonic narrative construction and appreciation has also been formulated as an ‘artistic research objective’. However, not all participants felt in command of this interaction mode.

Contrasting the works that emerged from the two installations is very interesting. The difference in the graphic works in the bottom of Figure 1 is striking in this respect. This difference could also be easily sensed when contrasting the music, dance, computer music pieces, and in the essays that were created in response to the two installations. It may be interpreted as a commentary on the differences in the

aesthetics and the user experience of interacting with each installation.

We believe that this limited account brings forward the ability of creative evaluation to stimulate thinking and the production of categorical statements but also commend on the aesthetics of each work.

Participants' experience with Creative Evaluation

We were pleasantly surprised by the ability of every single participant in the study to come up with imaginative content in response to the challenging tasks they were assigned. This was for us a sheer demonstration of the potential of human creativity.

Already during the first appointment, participants typically spent more than one hour interacting with the installations; visitors during public exposition rarely spent more than 20 minutes. We attribute the increased engagement and motivation to the fact that creative evaluation directly stimulated participant creativity as well as placed the responsibility for the evaluation outcome on the participants themselves. This is in contrast to the passive role participants often assume in user evaluation studies.

We believe there is merit in designing user evaluation to be a creative, stimulating, and active participatory experience. Participant creativity is valued and supported in other activities such as participatory design workshops for example. This has not been generally the case for user evaluation activities. Our study shows that there is space for incorporating creative practice within user evaluation activities. The increased participant engagement together with multidimensionality and richness of the evaluation outcome may greatly increase the chances of reaching user-centered conclusions when considering what would be worth revisiting in a specific work. In this sense, the incorporation of creative practice in evaluation may constitute a valuable alternative to task-oriented laboratory evaluation protocols.

Very important is to involve artists in the evaluation. This is because of their training in accessing non-propositional layers of experience, and in eliciting aesthetic thinking and knowledge.

Despite the unconstrained nature of the evaluation we used, converging points did appear in the outcome. This may indicate that creative evaluation protocols do not restrict the appearance of converging evidence. These may not only relate to what participants consider worth communicating, but also to the way this is presented in the evaluation outcome.

For example: Both musicians created pieces that used spatial analogies on their instruments. Both music aesthetics professionals used a free style in their essays¹⁰. Both dancers reinterpreted the movement repertoire they discovered while interacting with the installations to form their pieces. Both computer musicians created variations of the installations. These common aspects are surprising, especially given that the artists visited, planned, and executed their works alone.

We hardly expected any such consistencies when planning this study and are looking forward to a closer examination in the future.

Our experience with Creative Evaluation

The artistic works and the videos communicate emotion, provide an artistic and aesthetic view on the installations, and *show* us the installation experience. Difficulties, wishes, tensions, interpretations, and learning processes, they all can be observed in the videos in which participants enacted their experience and the artistic works. *Showing* rather than saying may have an important part to play in shaping evaluation techniques for interaction design.

A very interesting by-product of the process was the ability of the evaluation materials to stimulate thinking about the installations and allow us to come up with propositional feedback. In this sense, creative evaluation resulted in explicit, tacit, and ineffable enacted knowledge about the ways the installations may be experienced [2]. Their formulation using art is essentially an exercise in artistic research.

The participants in this evaluation experiment were artists. This was an intentional choice because of the ability of artists to formulate their work using aesthetics. Would creative evaluation be possible to apply if participants that do not have an artistic background were involved? After all, all participants delivered videos of their experience without being film professionals. In them, they enacted their experience, their learning, discoveries, and the way they constructed meaning by interacting with the installations. This material was useful for us, and we would like to think that people that have not been trained in the arts can potentially find ways to compose their experience in an artistic medium. The result is valuable because it contains what participants consider worth communicating in a rich non-verbal form. But we believe that it cannot be generally assumed that such outcomes will embody and communicate artistic thinking and knowledge. Our impression is that the skill associated with creating illustrative representations or composing artifacts that embody aesthetic thinking and knowledge should not be underestimated.

The evaluation results were well-received by the artists. Both agreed that the feedback from the videos and especially from the artistic works was novel and relevant to their practice. Importantly, their exposure to the evaluation outcome may be the first step towards an interdisciplinary, non-propositional dialogue. Other than an evaluation tool, such dialogues through art or design may turn out to be highly important for the exploration of design and artistic spaces and study topics. It remains to be seen how such approaches may combine with the traditional dialogue in HCI research that originates in the use of categorical feedback in propositional statements.

Some interesting observations relate to the public setting in which the final performances took place. This was not only because the open and public final event was a pleasant happening. Its value was significant both for the artists that participated in the evaluation and for us the researchers.

¹⁰despite being trained in the academic writing tradition

Although this event may appear, at a first glance, to be irrelevant for the evaluation, upon careful thinking and inspection it is only difficult to underestimate its importance.

In artistic practice, the public display of a work and its social dimension is part of the natural evolution of an artwork. Neglecting this, would be to deprive an artistic work of an essential driving force towards sharper form and clearer intentions. This risks relegating the work to the ‘preliminary study’ category. The evaluation outcome would therefore have been incomplete, had we deprived the artists from this opportunity.

Furthermore, being able to attend to the performances directly, unmediated by documentation, was also a primary aim for us. First, this shared experience provided an ecologically valid common ground for the evaluation. Simply working with the documented works would unavoidably influence our perception of them and risk obstructing conscious artistic choices. In addition, the live juxtaposition of the works revealed otherwise unimaginable connections of each work to the others.

Crucially, the *finissage* is an acknowledgement to the approach we have taken in dealing with the multi-dimensionality of aesthetic experiences. Instead of trying to “control” it, we made it the central theme of the evaluation and of the final event.

Limitations, clarifications, and future perspectives

We believe that there is potential in creative evaluation. Our initial experiment, although likely sufficient for introducing the concept, is only the beginning in the effort to experience how such techniques may be applied in interaction design.

This may be achieved by extended systematic application of the technique. A first step in this direction would be to use more artists of the same genre. This would allow the observation of within genre differences which may be especially revealing. Exposure to a greater variety of installations and extending the artistic practices involved would of course also help to understand what can be expected of such techniques.

Equally interesting would be to apply such evaluation techniques at earlier stages of the interactive system development. The installations we evaluated here were considered to be finished by their creators. Despite the limited space for further development in the context of these specific works, we are currently working with the artists in order to integrate the evaluation outcome in their praxis.

Applying similar evaluation experiments within design research, but also in closed task applications, may be very stimulating and revealing. An artistic commentary on contemporary interactive systems is much needed and would constitute an important contribution to user experience research. Contrasting the results of such evaluations with these of other evaluation techniques would allow us to better understand the way creative evaluation may be used to augment existing evaluation methods.

Importantly, it may also offer us new ways in understanding what is valuable within the context of interactive systems evaluation and contribute to the elucidation of the process of knowledge acquisition and communication through praxis [20, 25].

CONCLUSIONS

In our trying to overcome the difficulties with evaluating open-ended applications and to address the aesthetics of interactive systems, we turned our attention to the evaluation of interactive art. Reviewing the literature convinced us that traditional evaluation methods face difficulties in the absence of an ‘intended’ use or interpretation, and that the propositionality of contemporary evaluation techniques cannot address the aesthetics of interactive systems sufficiently. In response, we devised the concept of creative evaluation and experimented with its application. Ten artists were asked to use two sound installations as inspiration and both direct themselves enacting their own interaction experience in short film and express it using the artistic medium they were skilled at. The evaluation outcome stirred both us, the evaluators, and the artists that created the installations. It showed a significant potential in structuring evaluation by asking participants to creatively engage with the evaluated works. Furthermore, the use of artistic media in evaluation provided an important way for addressing artists at a level in which aesthetic thinking and knowledge can be demonstrated efficiently.

ACKNOWLEDGEMENTS

This work was supported by the Zukunftsfonds Steiermark Klangräume Project (PN:6067).

Alexander Deutinger (dance), Lorenzo Derinni (violin), David Buschmann (music aesthetics), Davide Gagliardi (computer music), and Patricia Wess (design) participated in the Random Access Lattice evaluation.

Margarethe Maierhofer-Lischka (double-bass), Danielle Sofer (music aesthetics), Kosmas Giannoutakis (computer music), Eva Stern (design), and Eva Brunner (dance) participated in the Paris Flâneur evaluation.

The authors would like to thank Nick Acorne for documenting the evaluation and its outcome. Special thanks go to Reni Hofmüller for providing us with the opportunity to work in the esc medien kunst labor.

The authors are indebted to Gerhard Eckel and Martin Rumori for providing us with their sound installations, support for setting up, helpful discussions on the evaluation, and perspectives on its outcome. Thanks to Seppo Gründler for his continuous support and guidance.

REFERENCES

1. Javier A Bargas-Avila and Kasper Hornbæk. 2011. Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2689–2698.

2. Michael Biggs. 2004. Learning from Experience: approaches to the experiential component of practice-based research. *Forskning, Reflektion, Utveckling* (2004).
3. Michael Biggs and Henrik Karlsson. 2010. *The Routledge companion to research in the arts*. Routledge.
4. Zafer Bilda. 2006. Evaluating audience experience with interactive art. (2006).
5. Z. Bilda, L. Candy, and E. Edmonds. 2007. An embodied cognition framework for interactive experience. *CoDesign* 3, 2 (June 2007), 123–137.
6. Zafer Bilda, Ernest Edmonds, and Linda Candy. 2008. Designing for creative engagement. *Design Studies* 29, 6 (2008), 525–540.
7. Kirsten Boehner, Phoebe Sengers, and Simeon Warner. 2008. Interfaces with the Ineffable: Meeting Aesthetic Experience on Its Own Terms. *ACM Trans. Comput.-Hum. Interact.* 15, 3, Article 12 (Dec. 2008), 29 pages.
8. Henk Borgdorff. 2011. *The Routledge Companion to Research in the Arts*. Routledge, Chapter The production of knowledge in artistic research.
9. Hendrik Anne Henk Borgdorff. 2012. *The conflict of the faculties: Perspectives on artistic research and academia*. Leiden University Press.
10. Linda Candy. 2011. Research and creative practice. *Interacting: art, research and the creative practitioner* (2011), 33–59.
11. Linda Candy and Ernest Edmonds. 2011. *Interacting: Art, Research and the Creative Practitioner*. Libri Pub.
12. Linda Candy and Sam Ferguson. 2014. *Interactive Experience in the Digital Age: Evaluating New Art Practice*. Springer.
13. Stuart K Card, Thomas P Moran, and Allen Newell. 1980. The keystroke-level model for user performance time with interactive systems. *Commun. ACM* 23, 7 (1980), 396–410.
14. John Dewey. 2005. *Art as experience*. Penguin.
15. Paul Dourish. 2004. What we talk about when we talk about context. *Personal and ubiquitous computing* 8, 1 (2004), 19–30.
16. Ernest Edmonds, Zafer Bilda, and Lizzie Muller. 2009. Artist, evaluator and curator: three viewpoints on interactive art, evaluation and audience experience. *Digital Creativity* 20, 3 (2009), 141–151.
17. Ernest Edmonds and Linda Candy. 2011. *Interacting: art, research and the creative practitioner*. Libri.
18. Christopher Frayling. 1993. Research in art and design. (1993).
19. William Gaver. 2007. Cultural commentators: Non-native interpretations as resources for polyphonic assessment. *International journal of human-computer studies* 65, 4 (2007), 292–305.
20. William Gaver. 2012. What should we expect from research through design?. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 937–946.
21. William W Gaver, Jacob Beaver, and Steve Benford. 2003. Ambiguity as a resource for design. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 233–240.
22. Steve Harrison, Phoebe Sengers, and Deborah Tatar. 2011. Making Epistemological Trouble: Third-paradigm HCI As Successor Science. *Interact. Comput.* 23, 5 (Sept. 2011), 385–392.
23. Steve Harrison, Deborah Tatar, and Phoebe Sengers. 2007. The three paradigms of HCI. In *Alt. Chi. Session at the SIGCHI Conference on Human Factors in Computing Systems San Jose, California, USA*. 1–18.
24. Jonathan Hook, Rachel Clarke, John McCarthy, Kate Anderson, Jane Dudman, and Peter Wright. 2015. Making the Invisible Visible: Design to Support the Documentation of Participatory Arts Experiences. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 2583–2592.
25. Kristina Höök and Jonas Löwgren. 2012. Strong concepts: Intermediate-level knowledge in interaction design research. *ACM Transactions on Computer-Human Interaction (TOCHI)* 19, 3 (2012), 23.
26. Kristina Höök, Phoebe Sengers, and Gerd Andersson. 2003. Sense and sensibility: evaluation and interactive art. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 241–248.
27. Kasper Hornbæk. 2006. Current practice in measuring usability: Challenges to usability studies and research. *International journal of human-computer studies* 64, 2 (2006), 79–102.
28. J Gary Knowles and Ardra L Cole. 2008. *Handbook of the arts in qualitative research: Perspectives, methodologies, examples, and issues*. Sage.
29. Celine Latulipe, Erin A Carroll, and Danielle Lottridge. 2011. Evaluating longitudinal projects combining technology with temporal arts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1835–1844.
30. Dieter Mersch. 2015. *Epistemologien des Ästhetischen*. Zürich: Diaphanes.
31. Ann J Morrison, Peta Mitchell, and Margot Brereton. 2007. The lens of ludic engagement: evaluating participation in interactive art installations. In *Proceedings of the 15th international conference on Multimedia*. ACM, 509–512.

32. Donald A Norman. 2013. *The design of everyday things: Revised and expanded edition*. Basic books.
33. Marianne Graves Petersen, Ole Sejer Iversen, Peter Gall Krogh, and Martin Ludvigsen. 2004. Aesthetic Interaction: a pragmatist's aesthetics of interactive systems. In *Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques*. ACM, 269–276.
34. Virpi Roto, Marianna Obrist, and Kaisa Väänänen-Vainio-Mattila. 2009. User experience evaluation methods in academic and industrial contexts. In *Interact 2009 conference, User Experience Evaluation Methods in Product Development (UXEM'09)*, Uppsala, Sweden. Citeseer.
35. Maggi Savin-Baden and Claire Howell Major. 2013. *Qualitative research: The essential guide to theory and practice*. Routledge.
36. Stephen Scrivener and others. 2002. The art object does not embody a form of knowledge. (2002).
37. Phoebe Sengers and Bill Gaver. 2006. Staying open to interpretation: engaging multiple meanings in design and evaluation. In *Proceedings of the 6th conference on Designing Interactive systems*. ACM, 99–108.
38. Richard Shusterman. 1992. *Pragmatist aesthetics: Living beauty, rethinking art*. Vol. 27. Cambridge Univ Press.
39. Peter Wright, Jayne Wallace, and John McCarthy. 2008. Aesthetics and Experience-centered Design. *ACM Trans. Comput.-Hum. Interact.* 15, 4, Article 18 (Dec. 2008), 21 pages.
40. James O Young. 2003. *Art and knowledge*. Routledge.