Dalia El-Shimy and Jeremy R. Cooperstock
Center for Interdisciplinary Research in Music, Media, and Technology
McGill University
527 Sherbrooke Street West
Montreal, Quebec H3A 1E3, Canada
{dalia, jer}@cim.mcgill.ca

User-Driven Techniques for the Design and Evaluation of New Musical Interfaces

Abstract: The merits of user-driven design have long been acknowledged in the field of human–computer interaction (HCI): Closely involving target users throughout the lifecyle of a project can vastly improve their experiences with the final system. Thus, it comes as no surprise that a growing number of music technology researchers are beginning to incorporate user-driven techniques into their work, particularly as a means of evaluating their designs from the perspectives of their intended users. Many, however, have faced the limitations that arise from applying the task-based, quantitative techniques typically encountered in classical HCI research to the evaluation of nonutilitarian applications. The nature of musical performance requires that designers reevaluate their definitions of user "goals," "tasks," and "needs." Furthermore, within the context of performance, the importance of creativity and enjoyment naturally supersedes that of efficiency, yet these concepts are more difficult to evaluate or quantify accurately.

To address these challenges, this article contributes a set of key principles for the user-driven design and evaluation of novel interactive musical systems, along with a survey of evaluation techniques offered by new directions in HCI, ludology, interactive arts, and social-science research. Our goal is to help lay the foundation for designers of new musical interfaces to begin developing and customizing their own methodologies for measuring, in a concrete and systematic fashion, those critical aspects of the user experience that are often considered too nebulous for assessment.

In 1986, Norman and Draper popularized the notion of "user-centered design," or the process of systematically involving users throughout a system's design and development cycles, with their seminal book on the topic, *The Design of Everyday Things*. To summarize the authors' view, user-centered design is the attempt "to ask what the goals and needs of the users are, what tools they need, what kind of tasks they wish to perform, and what methods they prefer to use" (Norman and Draper 1986, p. 2). Concurrently, Gould and Lewis (1985) devised a concrete user-centered methodology by distilling the best known "user-centric" practices from research into human–computer interaction (HCI) at the time. Their approach was based on three key principles: early focus on users and tasks, empirical measurement, and iterative design. Together, these principles aimed to provide a general guideline for designers, who were then advised to further define and choose the specifics of user involvement in their own work.

As user-centered design gained popularity in the ensuing years, it also became a topic of interest among designers of new musical interfaces (NMIs), a term by which we describe novel interactive music systems, gestural controllers, sound installations, and sonic environments. This can in part be attributed to the fact that music researchers had begun acknowledging the commonalities between their field and that of HCI research. Atau Tanaka (2000, p. 279), for instance, argues that instrumental music "establishes rich forms of human–machine interaction," and that, as a result, the successful design of musical interfaces should be "the result of a fusion of computer–human interface design and acoustic instrument lutherie" (Tanaka 2000, p. 403).

Although traditional user-centered or participatory design methodologies may prove suitable for the design of conventional computer-based systems, their applicability for creative or artistic interfaces has proved less clear, however. The nonutilitarian essence of musical performance poses special challenges, requiring attention to benchmarks, evaluation techniques, and alternatives to formal quantitative testing that are more suitable to its exacting and often subjective nature. Furthermore, user-driven evaluations of musical interfaces typically suffer from a lack of accepted or standardized frameworks (Geiger et al. 2008). Such difficulties are by no means unique to the context of musical

Computer Music Journal, 40:2, pp. 35–46, Summer 2016
doi:10.1162/COMJ_a_00357

performance, but are generally inherent to systems whose primary purpose is to support artistic expression or experience.

This article arose in the context of a project we undertook to explore the user-centered design of a novel distributed performance system, and during which we experienced the shortcomings of the principles articulated by Gould and Lewis when applied to the design of a musical system (El-Shimy and Cooperstock 2013; El-Shimy 2014). In turn, this prompted us to survey the collection of user-driven techniques that had proven to be better-suited for the design and evaluation of NMIs. Although we sought to focus on interfaces and computer-based instruments for musical creation and performance, we found that other disciplines, including ludology, interactive arts, affective computing, and the social sciences, all offered complementary tools for examining the more elusive aspects of the user experience, such as creativity, enjoyment, aesthetics, and engagement. Through our experience integrating and applying such tools to our own research, we were able to distill a set of principles for the user-centered design of NMIs. In turn, we endeavor not only to provide a review of the literature involving the use of these techniques and their respective benefits, but also to contribute guidelines for designers wishing to apply this large collection of tools to their own efforts.

Music-Oriented HCI

Tanaka (2006) explains that the design of NMIs "should benefit from techniques from human–computer interaction research." Such views have helped motivate the emergence of "music-oriented HCI" research, where the development of new sensing technologies, the creation of mapping strategies, and user involvement in design are heavily driven by HCI proficiency. Traditionally, much research in this area was devoted to using knowledge from HCI to match input/output paradigms suitably to musical tasks (Bongers 2000). Developers of NMIs soon began taking an interest in HCI research beyond mapping and interaction design, choosing instead to explore and adopt user-centric methodologies. A notable example is the work of Wanderley and Orio (2002, p. 69), who posit that "results from HCI can suggest methodologies for evaluating controllers, provided the context of interaction is well defined." Inspired by Bill Buxton's work on the assessment of input devices, the authors argue that the user-centric evaluation of novel input devices can best be accomplished when such devices are matched to potential applications using simple, representative musical tasks. These tasks, they add, should be designed to account for important parameters of usability within the context of musical performance, namely: "learnability," "explorability," feature controllability, and timing controllability (Orio, Schnell, and Wanderley 2001).

Although other researchers have found the techniques proposed by Wanderley and Orio useful to evaluate various musical tasks performed by using controllers, there were also limitations. In fact, even prior to Wanderley and Orio's work, Brad Cariou (1992, p. 366) had observed that "it is not only undesirable but impossible to define the musician's task." Furthermore, Kiefer, Collins, and Fitzpatrick (2008, p. 87) found they could not capture "in the moment" data about the user experience, something they believe to be important for musical evaluation, and they attribute this problem to the fact that "HCI methodology has evolved around a task-based paradigm and the stimulus-response interaction model of WIMP [windows, icons, menus, pointer] systems, as opposed to the richer and more complex interactions that occur between musicians and machines." In a similar vein, Johnston, Candy, and Edmonds (2008, p. 563) wrote that "software designed to facilitate musical expression presents a problem in this context, as it is difficult to formulate tasks to assign to users that are measurable but also meaningful." Finally, Stowell et al. (2009, p. 960) stated that "live music-making using interactive systems is not completely amenable to traditional HCI evaluation metrics such as task-completion rates."

As a result, the type of information that designers elicit from users, and the manner in which they elicit such information, have been topics of much discussion among creators of NMIs and digital arts who are keen on adopting user-centered design methodologies. Although researchers agree that user

involvement can provide much-needed structure to musical interface design, there is less of a consensus when it comes to deciding the exact nature of this involvement (Poepel 2005; Geiger et al. 2008). To assist designers of new musical interfaces in resolving these challenges, we present a number of key principles conceptually inspired by the approach that Gould and Lewis (1985) took within the context of HCI. And, as was the case with Gould and Lewis, our key principles are not meant to provide a rigid framework for designers, but to offer them general guidelines as they define the specifics of user involvement in their own work.

Key Principles

We propose three key principles for the user-driven evaluation of new musical interfaces, namely: (1) validate the basics, (2) investigate suitable alternatives to "usability," and (3) tailor evaluation techniques. Each of these principles is explored in detail in the following sections.

Validate the Basics

Developers looking to support or augment creative or artistic activities may encounter difficulties when defining system functionality that is to meet concrete user needs at the start of a project. This is especially the case if the goal is to introduce technology that is completely novel to the typical end user. In this section, we argue that developers may overcome this barrier by first developing an early and thorough understanding of basic user interactions that may be related to the technology under consideration.

Novel instruments are often created by artists to support their expression of a particular message or idea. As a result, many such instruments end up being used exclusively by their own designers. A growing number of researchers are, however, creating instruments meant for use by a wider audience. Accordingly, they are investigating the factors that must be addressed to encourage adoption by additional users. For instance, during his work on Faust Music On-Line, an interface designed specifically with accessibility to novice musicians in mind, Sergi Jordà (2002) identified idiosyncrasy as the biggest problem preventing new musical controllers from reaching a wider audience. In fact, Orio, Schnell, and Wanderley (2001) describe the entire design of NMIs as "marked by an idiosyncratic approach," especially when compared with the design of input devices in HCI. Cornelius Poepel (2005) attributes the problem of adoption in part to the fact that evaluation of NMIs is "often done by the developer or a small number of people." This phenomenon tends to occur because developers of NMIs typically see themselves as one and the same as their target users. Their understanding of every aspect of the system, however, prevents them from knowing what other users—even ones with the same level of musicianship—may perceive as complex. This problem may be exacerbated by the lack of established guidelines, missing interface standards, and a relative paucity of literature for the design and evaluation of NMIs, leading many developers to instead adopt a "trial-and-error" approach (Geiger et al. 2008; Stowell et al. 2009).

Morreale, Angeli, and O'Modhrain (2014) attempted to address such issues, in part, through their Musical Interfaces for User Experience Tracking (MINUET), a user-centered framework consisting of two stages: "goals," whereby designers define the desired user outcomes of a new musical interface, and "specifications," in which the appropriate user interactions are designed. Although this example of music-oriented HCI can help provide structure to the design of NMIs, we argue that the musical context nevertheless continues to pose challenges to the traditional user-centered techniques from which MINUET takes its inspiration. First, performers represent a unique type of user: Their "needs" can be difficult to establish, given that novel artistic tools typically do not exist to serve a concrete purpose in the same manner as utilitarian tools. Their "goals" when using such tools can also be too ambiguous to define, given that they have perhaps never considered alternatives to their traditional gear. Furthermore, as with any physically and mentally demanding activity, the nature of musical performance imposes strict constraints

on any interaction design. As a result, many considerations of usability design bear an added level of complexity, and many traditional input/output paradigms become unsuitable.

This implies that designers applying user-driven techniques must begin by thoroughly understanding how users undertake the most basic activities related in nature to the intended system. Such an approach is consistent with Gould and Lewis's first principle of early focus on users and tasks. This understanding can serve as the foundation for conceptualizing elementary system functionality, which in turn serves as a helpful starting point for development, as was the case in our experience.

Consider this: A challenge we faced while designing novel environments for real-time distributed performance was that our target users typically lacked prior experience with similar systems. As a result, we could not anticipate the types of interaction they would find useful, or the nature of the problems they might encounter in such a context. Similar to the approach adopted by Xambó and coworkers (2011), our initial prototype was designed to capitalize on familiar interactions and behaviors that would inform discussions with and elicit meaningful feedback from our target users. As our prototype evolved, we continued to refer to that feedback, ensuring that we respected the users' behaviors, needs, and interactions (El-Shimy 2014).

Similarly, and as part of a broader practice-based approach—a research process whereby new knowledge is gained by partaking in a certain type of practice and examining the outcomes—Andrew Johnston (2011) began the design of several NMIs by investigating the habits and behaviors of the musicians and composers involved in the project's lifecycle. Through online diaries, interviews, and software version control logs, he was able to derive design criteria against which to evaluate his prototypes, a process that helped further his understanding of musician–instrument interaction within the context of NMIs. Owen Green (2014) has also advocated a practice-led approach—a similar yet distinct form of research where the primary concern is understanding the nature of and furthering knowledge about a particular practice—as a means of engaging with questions of musicality. Although no specific methodology is offered, Green provides two case studies that serve to illustrate the notion that studying the practice of performance itself can serve to complement the more technical aspects of interaction design in music technology research.

Investigate Suitable Alternatives to "Usability"

Classical approaches to user-centered design have typically placed a strong emphasis on the degree to which users could successfully perform specific tasks with a given system under evaluation. With a growing number of varied disciplines turning to HCI research for guidance on designing not only usable but also engaging systems, however, many researchers were faced with the shortcomings of such a task-based approach. Liam Bannon (2005), for instance, calls for a better framework for conceptualizing human activities both at the interpersonal and behavioral levels. Similarly, Kaye et al. (2007, p. 2118) ask, "what of technology not for accomplishing tasks but for having experiences, for expressing one's identity, for flirting and arguing and living?" In turn, the inadequacies of the task-based approach in examining performance aspects beyond usability have led to the emergence of what is now known as "third-wave" or "third-paradigm" HCI, a trend described by Kiefer, Collins, and Fitzpatrick (2008, p. 89) as a "a response to the evolving ways in which technology is utilized as computing becomes more increasingly embedded in daily life." Third-wave HCI promotes an experience- rather than a task-based approach to user-driven design. It encourages what Fallman and Waterworth (2005, p. 1) describe as a focus on "experiences rather than performance; fun and playability rather than error rate; and sociability and affective qualities rather than learnability." As a result, third-wave HCI is particularly suited to the design and evaluation of novel interactive musical interfaces.

A parallel view within the musical context is presented by Johnston (2011, p. 280), who posits that "evaluation is best seen as a component of a broader examination of musical interface

design and musical expression," and as such argues for the more appropriate term "user-experience study." He further calls on music technology researchers to design, refine, and redefine the criteria according to which their novel instruments are developed and examined. Johnston's ideas are also reflected in the experience-based HCI approach we adopted during our user-driven design of novel systems for collaborative distributed performance (El-Shimy, Hermann, and Cooperstock 2012). As early as our user-observation stage, it became apparent that several aspects of traditional task-based usability testing would not be applicable to the system we had in mind: To what extent could we suitably define a user task within collaborative musical performance, particularly when the focus is on performer–performer rather than performer–instrument interactions? And if musicians did not strictly need to take advantage of our system's features, how could we evaluate their level of satisfaction? As a result, we sought to investigate, in as systematic a fashion as possible, which evaluation criteria would specifically fit musicians and their expectations. This was accomplished through early user interviews, to which we applied a qualitative data analysis (a technique we discuss later in this article) that uncovered such benchmarks as enjoyment, creativity, and self-expression as being of utmost importance to musicians.

As a result, we encourage designers interested in taking a user-driven approach to the design of nonutilitarian systems to investigate alternatives to the traditional notion of "usability." This is possible by uncovering and defining benchmarks specifically suited to the activity at hand. In turn, such benchmarks can lead to a more reliable evaluation of the system against the target user's expectations. To assist in this regard, in the following sections we explore what Bilda, Edmonds, and Candy (2008) describe as "fun, pleasure, goodness, and beauty": those facets of interaction considered critical to the user experience.

Affect

Typically, musicians and artists express a greater interest in the hedonic aspects of their experience with a system than they do in the system's efficiency or practicality. A number of researchers, however, have demonstrated that users of practical applications also exhibit a strong appreciation for other, less pragmatic qualities of interaction. In fact, Don Norman (2002) notes that emotions can change the way we approach a problem, making cognition and affect—processes that lead to understanding and evaluation, respectively—a "powerful team."

Although the importance of affect in the design of engaging systems is widely acknowledged, there is less of a consensus on how such a quality is best evaluated. For instance, Isbister et al. (2006, p. 315) explain that "evaluation of user affect is a domain that is not as well articulated and explored as is assessing whether a system is usable, or whether it actively increases work productivity." To this, Hassenzahl, Beu, and Burmester (2001, p. 7) add that "traditional usability engineering methods are not adequate for analyzing and evaluating hedonic quality and its complex interplay with usability and utility." As a result, a number of researchers have designed novel evaluation techniques that specifically introduce a greater a level of rigor to the study of the more subjective aspects of interaction. Examples include the repertory grid technique (Fallman and Waterworth 2005), the semantic differential (Fallman and Waterworth 2005), AttrakDiff (see http://attrakdiff.de), structured hierarchical interviewing for requirement analysis (SHIRA; cf. Hassenzahl, Beu, and Burmester 2001), the sensual evaluation instrument (SEI; see Isbister et al. 2006), the product emotion measure (PrEmo; cf. Desmet, Hekkert, and Jacobs 2000), and AMUSE (Chateau and Merisol 2005). We note that, although these techniques could potentially be applied within a musical context (with AttrakDiff, for instance, having recently been utilized by Poepel et al. [2014] throughout the evaluation of a gesture-based singing installation), Kiefer, Collins, and Fitzpatrick (2008, p. 90) warn that they first "need to be assessed specifically in terms of evaluation of musical experience as well as user experience." As such, we encourage developers of new musical interfaces to consider these examples for inspiration, or adapt them to fit their needs as they deem suitable, an approach we further discuss later in this article.

Fun, Pleasure, and Flow

Wessel and Wright (2002) argue that although getting started with computer-based instruments should be easy, continued development of expressivity is a key factor in the adoption of these instruments. Sidney Fels (2004) further expounded on this view, explaining that a "well-designed instrument" is one comprising an interface that is constrained and simple enough to allow a novice to make sounds easily, while also remaining sufficiently challenging for the experienced player to explore a path to virtuosity. A similar view is adopted by MacDonald, Byrne, and Carlton (2006), who argue that "musical activities must provide players and composers with continually demanding challenges in such a way as to keep the individual interested, stimulated, and in flow." Furthermore, Bryan-Kinns, Healey, and Leach (2007) explored the concept of "group flow" as part of their study of mutual engagement within the context of collaborative musical interfaces. This notion of "flow" was first formalized by Mihaly Csikszentmihalyi to denote a state of optimal experience marked by a feeling of energized focus, full immersion, and enjoyment (Hassenzahl, Beu, and Burmester 2001; Csikszentmihalyi 2009). A state of flow occurs when a task presents a level of challenge that is perfectly matched to a user's skill set, thereby precluding overstimulation on one end and boredom on the other. It is typically characterized by a feeling of energized focus, a sense of reward, a merging of action and consciousness, and—naturally—a notable level of pleasure. In fact, Csikszentmihalyi originally conceived of flow while investigating the concept of enjoyment itself. As Csikszentmihalyi (1975, p. 182) explains, "in this flow state, people experience a narrow field of intense concentration, they forget about personal problems, feel competent and in control, experience a sense of harmony and union with their surroundings, and lose their ordinary sense of time." Thus, flow can be considered a reliable indicator of pleasure and enjoyment.

A number of HCI researchers have also explored the notion of "fun" in relation to overall usability. For instance, Nolan Bushnell (1996, p. 31) noted that "whimsy and fun are often the precursors to powerful tools that are used later for more serious applications," and Hassenzahl, Beu, and Burmester (2001) describe "joy of use" as an important dimension of overall usability that designers must consider, if only for the humanistic view that "enjoyment is fundamental to life." Nonetheless, although pleasure is clearly observable, defining the metrics to assess fun can be a challenge. As, for instance, Charlotte Wiberg (2005, p. 1) explains, "we have so little knowledge about how traditional usability evaluation works in the context of fun and entertainment work, it is difficult to argue for new approaches."

Although the joy of use has received relatively limited attention in HCI, it has been widely examined within the study of gaming, a field otherwise known as ludology, where the study of fun, pleasure, and flow is considered a cornerstone of game design. Example frameworks include Marc LeBlanc's "Eight Kinds of Fun" (see http://8kindsoffun.com), the GameFlow model (Sweetser and Wyeth 2005), and—most notably—the gaming experience questionnaire (GEQ; see IJsselsteijn et al. 2008). Designed to quantify the assessment of flow even further, the GEQ offers a set of in-game and postgame questionnaires targeted towards evaluating competence, sensory and imaginative immersion, tension, challenge, and negative and positive affects, many of which can be adapted towards a wide array of activities, including musical performance (El-Shimy, Hermann, and Cooperstock 2012).

Creativity

Creativity has always been considered an essential to most, if not all, artistic endeavors, including musical performance. In recent years, however, creative engagement has come to be regarded as an important quality to consider when designing interfaces meant not only for artistic purposes, but also for utilitarian ones as well. For instance, Candy and Hori note that there is a growing demand for information technology tools that can better support the needs of "creative users," such as professional knowledge workers. These users are increasingly relying on computers to facilitate creative aspects of their work, and to meet their demands, the authors

explain that "an understanding of the nature of creative cognition as well as an evaluation of the tools that are used in the creative process is needed" (Candy and Hori 2003, p. 46).

Nonetheless, the evaluation of creativity continues to be an open problem. As Candy and Hori (2003, p. 52) explain, "a creative act is, by its very nature, neither predictable nor repeatable." Kiefer, Collins, and Fitzpatrick (2008, p. 89) add that "getting people to perform a precise task can be difficult, especially when you have creative people performing a creative task." Within the context of interactive art, Bilda, Edmonds, and Candy (2008, p. 525) expand on this idea, stating that "by its very nature, creative engagement with interactive art systems is as varied as the individual people who interact with it," and that it is, "therefore, quite difficult to predict." So, although these authors acknowledge that a user-centric approach can be highly beneficial to the design, evaluation, and improvement of systems that promote creative engagement, they also recommend approaching the selection or design of any evaluation methodologies with great care. In fact, throughout their work on the the "beta_space" project, an experimental environment at the Powerhouse Museum in Sydney where "the public can engage with the latest research in art and technology," Bilda, Edmonds, and Candy opted for a practice-led approach. A studio environment was re-created in a research setting, allowing them to examine artists, curators, audience members, and even the researchers themselves involved in one aspect or other of the creative process. Various techniques such as the "think aloud" method, where subjects are encourage to verbalize their thought process, and the "codiscovery" method, where subjects discuss their interactions naturally with one another, were combined with traditional interviews, questionnaires, and observations to give researchers a thorough understanding of the various facets of creative engagement.

Tailor Evaluation Techniques

According to MacDonald and Atwood (2013, p. 1969), "evaluation has been a dominant theme in HCI for decades, but it is far from being a solved problem." This is particularly evident within the context of NMIs, where the user experience with playful or creative interfaces is often marked by an idiosyncratic quality (Orio, Schnell, and Wanderley 2001; Jordà 2002). As a result, we argue that idiosyncrasy in the evaluation of such interfaces may perhaps be an unavoidable phenomenon, making the search for a "one size fits all" solution potentially futile. Instead, we propose that developers investigate the possibility of adapting existing techniques, an approach also advocated by Kiefer, Collins, and Fitzpatrick (2008), or devising new ones if necessary. Selection from existing techniques, such as those described throughout this article, may in turn be motivated by several factors, including the availability of necessary tools, the degree to which a technique's intended context matches the one under examination, and the level of modification required to adapt a technique from one application domain to another.

For instance, after defining our evaluation criteria of enjoyment, interaction with others, creativity, and self-expression, we found that, to the best of our knowledge, standardized methods for assessing these factors had yet to be established. The GEQ, which largely encompasses questions on flow and immersion, proved to be a suitable contender for the evaluation of enjoyment because of the breadth of behaviors it examined (IJsselsteijn et al. 2008). In fact, the general nature of the questions from the GEQ meant that relatively little modification was necessary to adapt it to the musical context. Its enduring popularity also inspired us to devise our own questionnaires to evaluate additional benchmarks such as self-expression and creativity. Furthermore, in an example of the mixed research paradigm, we supplemented those questionnaires with qualitative open-ended discussions and quantitative logged data, thereby increasing the depth of feedback we could elicit from our users. This led, in turn, to marked improvements in our prototypes (El-Shimy, Hermann, and Cooperstock 2012; El-Shimy 2014).

To further assist designers with selecting and adapting to their needs any of the evaluation techniques detailed throughout this article, we present two frameworks that have proven to be

particularly suited to the study of nonutilitarian systems: qualitative research and mixed research.

Qualitative Research

The exploratory nature of qualitative research renders it quite suitable for developing mental models of user interaction with systems that are completely novel, or that use new technologies that have yet to be fully understood or documented. As new musical interfaces are often designed to explore such novel forms of interaction—interaction that also tends to encompass hedonic qualities that may prove impossible to quantify—their evaluation can be particularly effective when undertaken from a qualitative standpoint.

As an example, Ilsar, Havryliv, and Johnston (2014) relied exclusively on qualitative data when evaluating new mappings for AirSticks, an electronic percussion instrument, by analyzing observations and interviews with musicians interacting with the system. Furthermore, Johnston, Candy, and Edmonds (2008) investigated novel software that musicians could use along with their traditional instruments to create a mix of computer-generated and acoustic sounds. When it came to testing their system, they realized that, because it was intended to encourage musical exploration and "disrupt habitual ways of thinking about music," it would not be amenable to characterization by quantitative measurement. As a result, the authors designed a qualitative study anchored in content analysis, and encouraged the seven professional musicians who interacted with their system to "think out loud." Content analysis, a popular methodology in the social sciences, allows researchers to derive information from nonnumerical data. At its core, this technique operates on the principle of grounded theory, or the notion that hypotheses are contained within and can be induced from data collected during an experiment (Glaser and Strauss 1967). This is in contrast with traditional (and typically quantitative) scientific research, which postulates that hypotheses should be clearly formed before an experiment. Although interpretation of verbal and behavioral data is subjective by nature, content analysis introduces a certain level of rigor to the process: It relies heavily on a procedure known as coding, during which *codes*, or tags with predefined meanings, are assigned to events in a data set, such as behaviors obtained from user observations, or quotes obtained from user interviews. Coding is typically applied in an iterative fashion, whereby codes deemed sufficiently similar are grouped and combined, until a smaller, relatively stable set of codes emerges. From this resulting set of codes, researchers can begin to understand and formalize user motivations, tendencies, and goals. Grounded theory was also at the core of the GEO Landscapes interactive art project, where Bilda, Bowman, and Edmonds (2008) filmed participants' interactions with the work and performed a content analysis to understand their reactions and preferences. Similarly, through extensive observations and interviews, we applied grounded theory in our own work towards understanding the motivations behind group musical performance (El-Shimy, Hermann, and Cooperstock 2012).

The qualitative experiment is one technique that has proven particularly well suited to examining nonquantitative hypotheses. According to Gerhard Kleining (1986), who originally coined the term, and as translated by Ravasio, Guttormsen-Schär, and Tscherter (2004, p. 3), a qualitative experiment is "the intervention with relation to a (social) subject which is executed following scientific rules and towards the exploration of the subject's structure. It is the explorative, heuristic form of an experiment." The qualitative experiment begins with theorizing the existence of relationships and processes that are not only difficult to quantify, but also can only be quantified after additional special treatment. Subsequently, variables deemed related to such relationships and processes are examined in rigorous experimental settings analogous to those used in quantitative experiments. Where the qualitative experiment differs from its quantitative counterpart however, is, in the nature of the data collected: Interviews, discussions, case studies, and diaries are some examples of the techniques commonly used to elicit user feedback in the qualitative experiment. Ideally, the results of the qualitative experiment should serve to develop hypotheses that can, in turn, be verified through quantitative studies. Thus,

both approaches can effectively complement each other, providing, as mixed research typically does, a more-complete picture of the subject matter under consideration.

Ravasio, Guttormsen-Schär, and Tscherter (2004, p. 22) advocate use of the qualitative experiment when one's goal is "to discover (rather than to verify) structures, procedures, processes, and their inter-dependencies, and when the setting should be as close as possible to real-life . . . but still requires a degree of controlled removal of context." Thus, such a technique may prove beneficial to the study of new musical interfaces, systems that often exist to serve functions and address needs that may not necessarily be completely defined or understood during the early stages of design. In that same paper, the authors also sought to promote the use of rigorous and procedural qualitative methods by proposing a formal framework for the qualitative experiment, consisting of six possible strategies for observing test-dependent variables in a qualitative setting: (1) separation or segmentation, (2) combination, (3) reduction or attenuation, (4) adjection or intensification, (5) substitution, and (6) transformation. We successfully applied a combination of two such strategies during our evaluation of the various system features used to augment our novel distributed performance environment (El-Shimy 2014).

Mixed Research

Although qualitative methods are particularly suited to the early, exploratory phases of research design, their benefits become even more apparent when effectively combined with quantitative techniques. The resulting framework, known as mixed research, is described by Johnson and Onwuegbuzie (2004, p. 17) as "the class of research where the researcher mixes or combines quantitative and qualitative research techniques, methods, approaches, concepts, or language into a single study." Typically, quantitative research encompasses deduction, confirmation, theory and hypothesis testing, explanation, prediction, standardized data collection, and statistical analysis, whereas qualitative research focuses on induction, discovery, exploration, theory and hypothesis, the researcher "as the primary 'instrument' of data collection," and qualitative analysis (p. 18). Thus, one ideal combination, for instance, involves using qualitative methods to develop hypotheses that can subsequently be tested via quantitative techniques. Another approach is to conduct qualitative interviews to provide additional meaning and context to quantitative experiment data. In the end, researchers are encouraged to mix approaches individually by considering the advantages and disadvantages of each in light of the subject matter at hand. To help with such decisions, one can refer to the extensive list of strengths and weaknesses of both qualitative and quantitative research provided by Johnson and Onwuegbuzie.

As an example within the context of music-oriented HCI, Pras and Guastavino (2011) utilized mixed techniques by combining Likert-scale questionnaires with open-ended interviews, as part of their extensive study of the interactions between musicians, record producers, and sound engineers in the studio. Similarly, Kiefer, Collins, and Fitzpatrick (2008) relied on mixed research in their user-centric evaluation of musical controllers. With a philosophy influenced by HCI research, the authors supplemented logged quantitative data with the qualitative analysis of user comments made during and after test sessions. In the end, both types of techniques helped paint a more accurate picture of overall user performance with new musical interfaces. In addition, Bryan-Kinns, Healey, and Leach (2007) evaluated mutual engagement with the Daisyphone, a collaborative musical tool, by augmenting qualitative assessments of the interaction patterns between participants with quantitative measurements and subjective participant responses. Such data were used not only for evaluation of how the various system features affected mutual engagement, but also helped the researchers understand how best to assess mutual engagement.

To address the inadequacies of task-based HCI evaluation within the context of musical interaction, Stowell et al. (2009, p. 971) introduced two complementary techniques, one qualitative and the other quantitative. Their discourse analysis method is meant to extract "a detailed reconstruction of users' conceptualization of a system" through the

methodological analysis of social structures in discourse. On the other hand, the quantitative "musical Turing test" aims to produce a quantitative result on whether an interactive experience provided by a computational system is similar to that provided by a human. According to the authors, although the former represents a strong social constructionist attitude in which "key categories and concepts are not predetermined but are considered an important outcome of the analysis," the latter derives useful numerical results, albeit at the expense of imposing a predetermined conceptual framework on the interaction.

Overall, mixed techniques provide an expansive approach that not only helps researchers overcome the inherent limitations of individual methodologies but also promotes collaboration across multiple disciplines.

Conclusion

As exemplified by the growing body of "music-oriented HCI" work, many music technologists agree that the design of new musical interfaces should benefit from the wealth of techniques offered by human–computer interaction research. Traditionally, much research in this area was devoted to using knowledge from HCI to match input/output paradigms suitably to musical tasks. An increasing number of NMI designers are, however, turning to user-centric techniques—another fundamental area of HCI—as a means of refining their work. Nonetheless, the adoption of such techniques to the context of musical performance has its share of challenges. The problem lies in large part in the objective, quantifiable nature of performance indicators typically examined in traditional HCI task-based system evaluations. In contrast, much of a musician's experience with a musical interface can often be dictated by qualities that are subjective in nature: pleasure, creativity, aesthetic enjoyment, and engagement—none of which can be quantified directly (Stowell et al. 2009).

As a result, we are witnessing a shift away from traditional, task-based, usability-driven design towards third-wave HCI, which promotes experience-based design and evaluation, particularly within creative and artistic contexts. As described here, the experience-based approach has become increasingly common among designers of interactive arts, musical interfaces, and playful systems keen on adopting what Liam Bannon (2005) describes as a "human-centered perspective." These researchers are utilizing quantitative and qualitative techniques, and even developing entirely new tools, in an effort to create experiences that can closely match their target users' needs and expectations.

There remains an apparent lack of established conventions when it comes to conducting systematic evaluations of NMIs, however. As Wanderley and Orio (2002) point out, the wealth of creativity seen in the design of novel controllers, environments, and interfaces is countered by the lack of commonly accepted evaluation methodologies. As is evident by the breadth of techniques described throughout this article, however, we argue that that no one-size-fits-all solution exists for the design and evaluation of new musical interfaces. Instead, we believe that performance with new musical interfaces is an inherently idiosyncratic experience. As such, we encourage designers to tailor existing evaluation techniques, such as those discussed here, to their own needs, or even devise new ones if necessary. In an effort to assist designers through such a process, we contributed three key principles, which draw from our own experience, as well as existing literature on this topic: validate the basics, investigate suitable alternatives to usability, and tailor evaluation techniques. Just like Gould and Lewis's principles for usability, our principles are not intended to be rigid rules. Instead, they were developed to inspire and guide those wishing to adopt a user-driven approach to the design of novel musical, creative, or nonutilitarian systems.

References

Bannon, L. J. 2005. "A Human-Centred Perspective on Interaction Design." In A. Pirhonen, et al., eds. *Future Interaction Design*. London: Springer, pp. 31–51.

Bilda, Z., C. Bowman, and E. Edmonds. 2008. "Experience Evaluation of Interactive Art: Study of GEO

Landscapes." In *Proceedings of the Australasian Conference on Interactive Entertainment*, pp. 1:1–1:10.
Bilda, Z., E. Edmonds, and L. Candy. 2008. "Designing for Creative Engagement." *Design Studies* 29(6):525–540.
Bongers, B. 2000. "Physical Interfaces in the Electronic Arts: Interaction Theory and Interfacing Techniques for Real-Time Performance." In M. M. Wanderley and M. Battier, eds. *Trends in Gestural Control of Music*. Paris: IRCAM, pp. 41–70.
Bryan-Kinns, N., P. G. T. Healey, and J. Leach. 2007. "Exploring Mutual Engagement in Creative Collaborations." In *Proceedings of the ACM SIGCHI Conference on Creativity and Cognition*, pp. 223–232.
Bushnell, N. 1996. "Relationships between Fun and the Computer Business." *Communications of the ACM* 39(8):31–37.
Candy, L., and K. Hori. 2003. "The Digital Muse: HCI in Support of Creativity." *Interactions* 10(4):44–54.
Cariou, B. 1992. "Design of an Alternate Controller from an Industrial Design Perspective." In *Proceedings of the International Computer Music Conference*, pp. 366–367.
Chateau, N., and M. Merisol. 2005. "AMUSE: A Tool for Evaluating Affective Interfaces." In *CHI Workshop on Evaluating Affective Interfaces*. Available online at www.sics.se/~kia/evaluating_affective_interfaces/Chateau.pdf. Accessed 17 January 2016.
Csikszentmihalyi, M. 2009. *Flow*. New York: HarperCollins.
Desmet, P. M. A., P. Hekkert, and J. J. Jacobs. 2000. "When a Car Makes You Smile: Development and Application of an Instrument to Measure Product Emotions." *Advances in Consumer Research* 27:111–117.
El-Shimy, D. 2014. "Exploring User-Driven Techniques for the Design of New Musical Interfaces through the Responsive Environment for Distributed Performance." PhD dissertation, McGill University, Montreal, Canada.
El-Shimy, D., and J. R. Cooperstock. 2013. "Reactive Environment for Network Music Performance." In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 158–163.
El-Shimy, D., T. Hermann, and J. R. Cooperstock. 2012. "A Reactive Environment for Dynamic Volume Control." In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 378–383.
Fallman, D., and J. Waterworth. 2005. "Dealing with User Experience and Affective Evaluation in HCI Design: A Repertory Grid Approach." In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pp. 2–7.
Fels, S. 2004. "Designing for Intimacy: Creating New Interfaces for Musical Expression." *Proceedings of the IEEE* 92(4):672–685.
Geiger, C., et al. 2008. "Towards Participatory Design and Evaluation of Theremin-Based Musical Interfaces." In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 303–306.
Glaser, B., and A. L. Strauss. 1967. *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago, Illinois: Aldine.
Gould, J. D., and C. Lewis. 1985. "Designing for Usability: Key Principles and What Designers Think." *Communications of the ACM* 28(3):300–311.
Green, O. 2014. "NIME, Musicality, and Practice-Led Methods." In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 1–6.
Hassenzahl, M., A. Beu, and M. Burmester. 2001. "Engineering Joy." *IEEE Software* 18(1):70–76.
IJsselsteijn, W., et al. 2008. "Measuring the Experience of Digital Game Enjoyment." In *International Conference on Methods and Techniques in Behavioral Research*, pp. 88–89.
Ilsar, A., M. Havryliv, and A. Johnston. 2014. "Evaluating the Performance of a New Gestural Instrument Within an Ensemble." In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 339–342.
Isbister, K., et al. 2006. "The Sensual Evaluation Instrument: Developing an Affective Evaluation Tool." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1163–1172.
Johnson, R. B., and A. J. Onwuegbuzie. 2004. "Mixed Methods Research: A Research Paradigm Whose Time Has Come." *Educational Researcher* 33(7):14–26.
Johnston, A. 2011. "Beyond Evaluation: Linking Practice and Theory in New Musical Interface Design." In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 280–283.
Johnston, A., L. Candy, and E. Edmonds. 2008. "Designing and Evaluating Virtual Musical Instruments: Facilitating Conversational User Interaction." *Design Studies* 29(6):556–571.
Jordà, S. 2002. "FMOL: Toward User-Friendly, Sophisticated New Musical Instruments." *Computer Music Journal* 26(3):23–39.
Kaye, J., et al. 2007. "Evaluating Experience-Focused HCI." In *CHI Extended Abstracts on Human Factors in Computing Systems*, pp. 2117–2120.
Kiefer, C., N. Collins, and G. Fitzpatrick. 2008. "HCI Methodology for Evaluating Musical Controllers: A

Case Study." In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 87–90.
Kleining, G. 1986. "Das qualitative Experiment." *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 38(4):724–750.
MacDonald, C. M., and M. E. Atwood. 2013. "Changing Perspectives on Evaluation in HCI: Past, Present, and Future." In *CHI Extended Abstracts on Human Factors in Computing Systems*, pp. 1969–1978.
MacDonald, R., C. Byrne, and L. Carlton. 2006. "Creativity and Flow in Musical Composition: An Empirical Investigation." *Psychology of Music* 34(3):292–306.
Morreale, F., A. D. Angeli, and S. O'Modhrain. 2014. "Musical Interface Design: An Experience-Oriented Framework." In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 467–472.
Norman, D. A. 2002. *The Design of Everyday Things*. New York: Basic.
Norman, D. A., and S. W. Draper. 1986. *User Centered System Design: New Perspectives on Human–Computer Interaction*. Hillsdale, New Jersey: L. Erlbaum.
Orio, N., N. Schnell, and M. M. Wanderley. 2001. "Input Devices for Musical Expression: Borrowing Tools from HCI." In *Proceedings of International Conference on New Interfaces for Musical Expression*, pp. 467–472.
Poepel, C. 2005. "On Interface Expressivity: A Player-based Study." In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 228–231.
Poepel, C., et al. 2014. "Design and Evaluation of a Gesture Controlled Singing Voice Installation." In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 359–362.
Pras, A., and C. Guastavino. 2011. "The Role of Music Producers and Sound Engineers in the Current Recording Context, as Perceived by Young Professionals." *Musicae Scientiae* 15(1):73–95.
Ravasio, P., S. Guttormsen-Schär, and V. Tscherter. 2004. "The Qualitative Experiment in HCI: Definition, Occurrences, Value, and Use." Available online at shirahime.ch/WhitePrincess/wp-content/uploads/2011/02/QualExp.pdf. Accessed 17 January 2016.
Stowell, D., et al. 2009. "Evaluation of Live Human–Computer Music-Making: Quantitative and Qualitative Approaches." *International Journal of Human–Computer Interaction Studies* 67(11):960–975.
Sweetser, P., and P. Wyeth. 2005. "GameFlow: A Model for Evaluating Player Enjoyment in Games." *ACM Computers in Entertainment* 3(3):3.
Tanaka, A. 2000. "Musical Performance Practice on Sensor-Based Instruments." In M. M. Wanderley and M. Battier, eds. *Trends in Gestural Control of Music*. Paris: IRCAM, pp. 389–405.
Tanaka, A. 2006. "Interaction, Experience and the Future of Music." In K. O'Hara and B. Brown, eds. *Consuming Music Together*. Berlin: Springer, pp. 267–288.
Wanderley, M. M., and N. Orio. 2002. "Evaluation of Input Devices for Musical Expression: Borrowing Tools from HCI." *Computer Music Journal* 26(3):62–76.
Wessel, D., and M. Wright. 2002. "Problems and Prospects for Intimate Musical Control of Computers." *Computer Music Journal* 26(3):11–22.
Wiberg, C. 2005. "Affective Computing versus Usability? Insights of Using Traditional Usability Evaluation Methods." Paper presented at the CHI 2005 Workshop on Innovative Approaches to Evaluating Affective Interfaces, 4 April 2005, Portland, Oregon. Available online at www8.informatik.umu.se/~colsson/articles/CHI2005workshop.pdf. Accessed 17 January 2016.
Xambó, A., et al. 2011. "Collaborative Music Interaction on Tabletops: An HCI Approach." In *BCS HCI Workshop on When Words Fail: What Can Music Interaction Tell Us about HCI?* Available online at oro.open.ac.uk/29115/1/xambo_BHCI-2011-position-paper-CR-final.pdf. Accessed 17 January 2016.