Evaluating Improvisational Interfaces

Toby Gifford
SensiLab, Monash University
Melbourne, Australia
toby.gifford@monash.edu

Shelly Knotts
SensiLab, Monash University
Melbourne, Australia
shelly.knotts@monash.edu

Stefano Kalonaris
SARC, Queens University
Belfast, UK
stefanokalonaris@gmail.com

Jon McCormack
SensiLab, Monash University
Melbourne, Australia
jon.mccormack@monash.edu

ABSTRACT
Improvisational interfaces are human-machine systems for improvising creatively with. To facilitate research in this field, we seek some general approaches to evaluating the quality of a particular interface, and the success of a given improvisation. Drawing on a broad array of improvisational practices across music, visual art and drama, we draw out common threads and discuss the potential for implementing computational models of improvisation that can self-evaluate.

ACM Classification Keywords
Human-Centred Computing: HCI design and evaluation methods

Author Keywords
Improvisation; Computational Creativity; Evaluation

INTRODUCTION
Improvisation is well established mode of creative practice across many different domains: from music, through painting to drama, performance and beyond. Traditionally, improvisation research has mostly considered improvisation between human artists. However, in recent years, interest in improvising with non-human intelligences or systems has grown. We feel that the field has reached a level of maturity that now requires detailed understanding of both the mechanics and psychology of improvisation with non-human partners.

The focus of this paper is in improvisation with computer systems. It is part of a research trajectory in designing, implementing and evaluating improvisational interfaces – computer interfaces for improvising in creative arts domains, as outlined in [33]. Ultimately we are interested in realising the notion of machines as creative collaborators, and as part of this overall goal we seek to implement improvisational behaviours in machine agents, facilitate improvised exchanges between humans and computers, and create guidelines for designing interfaces that support improvised interaction.

As with any design research methodology, we see evaluation as playing a fundamental role [16]. A natural question presents itself: how does one evaluate an improvisational interface? For that matter, how does one evaluate an improvisation?

The design and implementation of machine improvisers and improvisational interfaces more generally is a trans-disciplinary endeavor, drawing from several fields, including psychology, computational creativity, and digital musical instrument (DMI) design.

The question of evaluation has been addressed by several authors in computational creativity (for example [26, 39]), in DMI design (for example [25, 4, 11]), and in the intersection of these two fields [43]. A common thread in these discussions is the importance of aesthetic evaluation of creative artefacts. It behoves us then to examine the critical cultures of various creative fields, and the kinds of aesthetic elements that enter into critical evaluations.

EVALUATING IMPROVISATION IN CREATIVE DOMAINS
Creative Improvisation
Many creative domains explicitly involve improvisation, particularly in performance: musical improvisations in various genres such as jazz, experimental, solos in rock; theater improvisation (in for example theater sports); spoken word in poetry slams and hip-hop battles.

Beyond performance, improvisation can play an important role in creative development for many artforms, such as music composition, painting, writing etc. Indeed Sawyer argues that improvisation is the “purest” form of creativity [41], the generative kernel of an iterative process of generation, evaluation and refinement, as depicted in figure 1.

Evaluation as Part of the Creative Process
According to Dewey, “creative activity is our great need, but criticism, self-criticism is the road to its release” [13]. In iterative models of the creative process evaluation plays a key role.
role in the iterative cycle, and some researchers argue that the creator is in a constant process of evaluation [35, 22, 41].

How, then, to evaluate an improvisation? As Derek Bailey wryly noted, “any subject is, I suppose, fair game for the academic, but it does seem that if there is anything which is singularly inappropriate for treatment in this way it must be free improvisation” [3]. Yet improvisation in any creative culture has ‘rules’ be they overt or covert.

**Anything Goes?**

A recurring theme in commentary from practitioners and theorists alike is countering the myth that in improvisation “anything goes”. For example:

> Improvisation is *not*, as it is sometimes perceived, completely free, or “doing whatever you feel like,” although it strives for new and less restricted avenues of expression than the current set of rules can accommodate – a process often involving intuitive and even rash decisions. It requires skill and training, it can be learned, and it can fail horribly, precisely *because* there are tacit rules within the community of improvisers [45].

Similarly, in reference to Jazz improvisation, Sawyer notes that “even the freest improviser plays within a musical tradition, and before you can improvise you have to learn that tradition” [41, p. 337].

**No Such Thing as a Mistake?**

“If you make a mistake then play the same thing over again and with feeling. Let them know that you meant it” (Louis Armstrong) [32].

Thomas Nunn seeks to dispel myths about absence of mistakes or difficulty in musical improvisation [36], and to, instead, approach improvisation as an expression that needs cultivating and practice, whereby mistakes can be contextualised and retrospectively subverted for incorporation into subsequent improvisations.

This process of recontextualising the unintended lets us consider that in improvised performance, evaluation is performed *prospectively* rather than retrospectively. The important question is not ‘how good was that?’, but rather ‘what can I do with that?’.

**Cultures of Critique**

One important methodological consideration that has received much support from researchers is the use of subjective aesthetic assessment by ‘domain insiders’ as an instrument for evaluating improvisation [43, 37, 30]. This parallels similar arguments in creativity assessment theory [1, 17].

In designing improvisational interfaces, it may be that the designer themselves play the role of expert judge. For example, in the context of designing DMIs, Jordá and Mealla comment:

> [Just as] much research in HCI culminates in lists of guidelines and/or principles for design (and/or evaluation of design) based on research or practical experience relating to how people learn and work, it comes as no surprise that the first tentative NIME design frameworks have been mostly proposed by experienced digital luthiers [25].

Ben Swift argues a similar line, suggesting that the term ‘criticism’ may be more apt than ‘evaluation’ in the context of DMI design:

> The concept of criticism, rather than evaluation, also provides an explicit scope for expert judgments. As Bardzell notes, expert judgments happen all the time in design, whether implicitly or explicitly. This has always been true for the design of musical instruments . . . we must not be afraid to take the same methodological stance in the design of [digital musical instruments] [44].

Embracing criticism as evaluation promises rich returns, since all creative domains have some culture of critique, and corresponding critical values. The difficulty of making explicit those values remains however. As Eisenberg and Thompson note, “the evaluation of improvised music is particularly mysterious” [17].

**ASPECTS OF IMPROVISATION**

Caines & Heble (2015) have compiled a collection of essays from improvisation practitioners and theorists of many persuasions. Reflective of the broad themes arising from the diverse contributions to the topic, they curatorially categorise these essays into seven aspects of improvisation, which also provides a point of departure for discussing the broad church of improvisation – listening, trust/risk, flow, dissonance, responsibility, liveness, surprise. We have chosen to variously conflate, modify, and extend their categories into the list: listening & liveness; trust, risk & responsibility; flow and rupture; and novelty & surprise.
**Trust, Risk and Responsibility**

From Caines and Heble’s more extensive list we’ve grouped together Trust, Risk and Responsibility as relating to the performer’s interaction with the social and physical constraints of improvisation. Improvisation fundamentally involves live exchange with audience, collaborators and interfaces, and the dynamics and boundaries of those interactions define the creative output. Trust, Risk and Responsibility are some of the parameters that define that interaction, and enable the performer to perform at all.

Trust, particularly, can be viewed as a heuristic decision-making construct which helps the parties involved in deliberating under conditions of uncertainty, which could potentially lead to inaction or paralysis. In this sense, trust is a construct/mechanism by which social complexity is reduced and actions made possible. In a musical improvisation context, trust could express the confidence one player has with respect to appropriate musical response and responsibility on behalf of the other player(s). Trust thus expresses the willingness of one party to rely and accept the actions of another party, often relinquishing control over the outcome.

Trust operates across many timescales of interaction: a virtuoso musician spends years developing an in-depth knowledge of their instrument developing technique and muscle memory. The improviser expects in performance that they will get a predictable output from a given input to the interface, and their trust in this allows them to concentrate on the creative side of the performance and to take calculated risks in performance. In a collaborative context, we can refer to “yes, and” rule of theater improvisation [14]. Performers trust that their collaborators will respond in a positive and creatively synergistic way to their contributions to a performance. One of the challenges of building computational improvisation partners has been developing this trust in given inputs generating predictable, but creative outputs.

These multiple layers of trust are important in allowing an improviser to take calculated risks in performance. We’ll explore risk here through a case-study from current practice in improvised computational arts. Risk is clearly exemplified in the growing field of Live Coding - a practice where performers write and modify their algorithms on stage to produce sound, visuals and other creative outputs. Live coders not only risk creative embarrassment, but risky programmatic moves may produce catastrophic system crashes or punctuate the performance with lengthy debugging processes. Live coders project their performance interfaces, giving the audience insight into their creative and computational processes, but also access to their computational failures. Culturally however this risk-taking is seen as part of the art-form. Anecdotally audience members have commented that crashing is ‘just what happens’ in live coding, and in more intimate settings ‘community bug fixing’ can come part of the performance narrative. Whereas in other computational arts rigorous testing is required to ensure system stability, live coders trust in their algorithmic knowledge to code their way out of creatively sticky corners and technical proficiency to avoid and solve errors. Arguably, risk is vital to generating novel creative outputs in performance - performers may combine algorithms in new ways and push the envelope of known outputs through complex algorithmic design.

Strongly related to trust and risk, responsibility may refer to the social responsibility to communicate with and respond to collaborators. Social interplay is key to successful collaboration, and improvisers continuously change roles during a performance, switching between leading, following, and taking part in duets, trios, etc. The dynamic nature of role assignment brings with it the necessity for performers to constantly by conscious of providing space for other performers but also taking the lead and contributing new ideas when creatively necessary. In computational creative domains we might also add to the social responsibility of the performer, the social and cultural responsibility of the technology designer [31, 27]. Beyond designing stable and reliable tools, the designer should also be conscious of how the affordances of their interfaces may impact on the socio-cultural factors of improvisation, and performers with appropriate levels of creative agency in performance.

**Liveness and Listening**

Improvisation is about the here and now – it is essentially situated in character [23]. The ability to respond to changing or unexpected circumstances – this is the hallmark of improvisation, but it is also the hallmark also of ‘liveness’.

What is it about live performance that thrills so, audience and performer alike? This has been the subject of much discussion in many art-forms, particularly music and theatre/film. In the context of digital media, prominent theses are those of Auslander [2] who, channelling Walter Benjiman, argues that live performance nows serves the role of ‘authenticating’ recorded works, and Dixon [15] who finds performance flourishing in technologically mediated artforms. Common threads include the demonstration of virtuosity and the concomitant necessity for risk-of-failure.

A common refrain amongst jazz musicians is the importance of being responsive when improvising [33, 5]. To improvise one must be aware of context, and synergistically interacting with other improvisers.

A central element of the phenomenon of peak or elevated performance in jazz improvisation is the placement in time on the part of the player. By this I do not mean rhythmic time, but rather a quite special sense, or ‘tri-focal’ awareness, of precisely (i) where one is in the heat of the creative moment, but with a simultaneous comprehension of (ii) where one in the context of the ensemble has come from and (iii) where one (and the ensemble) is going [23].

What of solo improvisers? Is this same sensitivity to context then important? In this case it becomes the audience with whom the improviser’s social contract is formed. In this sense the solo improviser is akin to any live performer. The live performer must construct the performance anew each time – conscious of the mood of the crowd, the feel of the place. The DJ, though often maligned by traditionalists as not really
performing, is fundamentally focused on this key facet of liveness.

Some of the authors’ previous experiments in machine improvisation have suggested that an awareness of, and responsiveness to, the actions of a human improvising partner, are among the most important traits of a machine improviser – in terms of fostering a sense of creative partnership [10]. Furthermore, in terms of information architecture, aesthetic evaluation performed by the machine improviser should be applied to the entire ensemble’s output as a whole (rather than, for example, simply the machine’s own output) [20, 21].

**Flow and Rupture**
An important phenomenological facet of improvisation is the experience of getting ‘in the zone’. Hagberg notes that Jazz improvisation creates in performers an “unmistakable sense of the inimitable uniqueness and non-arbitrariness of a performance in the zone – a state that has remained, despite its undeniable importance and special value in musical performance, interestingly easier to recognize than to fully understand and describe” [23].

**Cognitive Flow**
Reports, such as above, of a sense of focus and timelessness in performance have been defined by Czikzentmihalyi as cognitive Flow. This unhindered creative state can generally only be achieved when a performer is able to interact with their tools unobstructed by technical or physical inefficiencies, and crucially when various factors align to allow the performer to enter the correct cognitive state to be creative.

**Flow as ease of interaction**
Related to the experience of cognitive flow is the ability of the performer to have smooth and fluid interactions with an creative interface. Musical instruments have been refined over centuries such that the ergonomics of the interface allow a skilled musician to interact effortlessly with it to produce target sounds.

In designing new interfaces those that allow immediate, straightforward and unobstructed interaction with a creative output may have more appeal than those that are difficult to play or use. However, as McDermott et al. [34] argue, ‘ease’ in the typical HCI sense is not necessarily desirable in an improvisational interface. As discussed above, in order to obtain Flow, there should be a balance of skill and challenge – an interface that presents no challenges also does not afford demonstration of virtuosity in performance.

**Flow and Rupture as generative principles**
Another description of flow in relation to improvisation is given by Rose from the point of view of an observer. In breakdancing Rose describes the smooth transitions made by the performer from one point to the next. “Flow stands for the possibility while composing that each next move, whether abrupt or continuous, could present itself without hesitation or interruption.” [18].

Rose formulates this observance of flow as a generative device whereby new material is produced out of the performer finding a new state which can be smoothly transitioned to from the last state. Rose’s generative formulation relies on the idea of rupture as the antithesis to flow. Rupture is a point of stopping, changing direction or friction in the movement.

In relation to our current context on continuous evaluation, we could see rupture as reflective points where a performer breaks out of flow - generative, cognitive or interactive - to evaluate the performance and current context and decide on their next move. In this sense, aspects of friction and resistance may be just as important as flow in improvisation contexts as building opportunities for rupture into a system, interface or performance forces a performer to pause and reflect.

**Detecting and Evaluating flow**
The presence of flow is usually determined through various subjective observational factors evaluated after a creative activity such as experiencing timelessness, clarity and high concentration levels [24]. However, detectable physiological markers such as heart rate and breath depth have also been found to correlate with subjective observance of flow [12].

Performer feedback alongside physiological sensing could be useful in determining relative levels and occurrence rates of flow and indicating rupture frequency. This may allow an interface designer to determine the right level of friction into an interface and to determine the creative value of system resistance. Though physiological evaluation is likely only practical on a performance time scale [37], a diary of flow observance could apply across all time scales of interaction with an interface.

**Novelty and Surprise**
In simple terms, novelty is something either you, the world or the universe haven’t experienced before. More formally, novelty has been well studied in the creativity literature (see e.g. [7, 42]) where discussions about the relevancy of appropriate novelty play out (the qualitative difference between a statistical outlier and something perceived as genuinely surprising and new). More recently the concept of “novelty search” [29] has found application in evolutionary search. In this technique – inspired by the non-objectivity of real biological evolution the algorithm searches for regions of the search space not visited previously, deliberately avoiding the traditional fitness optimisation of standard evolutionary search. While mainly suited to deceptive problems, the idea has been further extended using “minimal criteria” [28] to reduce the overall size of the viable behaviour space and to try to minimise success of unsuccessful phenotypes.

Novelty and surprise are intrinsically linked to expectation. ‘When a listener (or performer as listener) hears what they expect, there is low complexity, and when they hear something unexpected, there is a higher complexity.’[8, p. 4] Therefore, one could say that low complexity, or low entropy in information-theoretical terms, violates surprise. The local importance of a particular musical event might vanish over time but, if subsequent material seems to relate to it, then its relevance will increase, prompting its inclusion in some schema which might help the forming of expectation regarding future events.
Some computer simulations of musical expectations have been used for computational music analysis [38] and for generative music production [9]. More broadly, computational evaluation of the balance of novelty and surprise has been part of several machine improvisation systems [6, 40, 19].

CONCLUSION

In order to progress research into improvisational interfaces, a set of guidelines, critical values, and methodologies for evaluating both interfaces and improvisations is desirable. Developing these is complicated by the diverse array of creative practices that involve improvisation, the variety of contexts in which such interfaces may be used, and the difficulty in formalising criteria for what makes a successful improvisation in any field. Nevertheless, every creative practice has a culture of criticism, and several common threads can be drawn from these that suggest general properties of successful improvisations, and in turn suggest methodologies for evaluating the efficacy of a given interface in facilitating improvisation.

Notions of trust, risk and responsibility are endemic in improvisation. One important aspect of implementing a computational collaborator is the handing over of responsibility for some part of the music production to the computer, with concomitant relinquishment of some degree of control. This requires trusting the machine to do something appropriate and interesting. Any useful evaluation of an improvisational interface will likely need to address this dimension, in terms of successful risk taking on the part of the machine, as well as the facilitation of successful risk taking on the part of the human.

Achievement of Flow is an important indicator of successful improvisations. Thus evaluating the attainment of flow by the performer is one approach to evaluating the interface with which they perform. Binary assessments of whether in-flow or not are readily self-reported by improvisers after the fact. More fine-grained measures of level-of-flow may also be gleaned from a variety of physiological markers.

A sense of immediacy is another key factor to evaluate in improvisational interfaces. The inherent temporality of improvisation implies that improvisational interfaces need to support immediate expression. This in turn implies the need for a rich input interface, whether by virtue of fluid gestural interfaces, or by leveraging rich physical systems such as acoustics or image processing.

Finally, the field of improvisation is broad and crosses many creative domains and creative cultures. As such no single evaluative framework is likely to fit all circumstances. Yet one methodological maxim seems to be universally applicable - ultimately improvisational interfaces should be evaluated through use, according to the quality of the interaction they engender, and their utility in furthering and facilitating creativity over many timescales, from a moment of performance to a life of learning.

ACKNOWLEDGMENTS

This research is supported by the Australian Government through the Australian Research Council’s Discovery Projects funding scheme (project DP160100166). The views expressed herein are those of the authors and are not necessarily those of the Australian Government or Australian Research Council.

REFERENCES


