Towards a Framework of Aesthetics in Sonic Interaction

Stuart Cunningham
s.cunningham@chester.ac.uk
University of Chester
Chester, UK

Iain McGregor
I.McGregor@napier.ac.uk
Edinburgh Napier University
Edinburgh, UK

Jonathan Weinel
J.R.Weinel@greenwich.ac.uk
University of Greenwich
London, UK

John Darby
j.darby@mmu.ac.uk
Manchester Metropolitan University
Manchester, UK

Tony Stockman
t.stockman@qmul.ac.uk
Queen Mary University of London
London, UK

ABSTRACT

As interaction design has advanced, increased attention has been directed to the role that aesthetics play in shaping factors of user experience. Historically stemming from philosophy and the arts, aesthetics in interaction design has gravitated towards visual aspects of interface design thus far, with sonic aesthetics being underrepresented. This article defines and describes key dimensions of sonic aesthetics by drawing upon the literature and the authors' experiences as practitioners and researchers. A framework is presented for discussion and evaluation, which incorporates aspects of classical and expressive aesthetics. These aspects of aesthetics are linked to low-level audio features, contextual factors, and user-centred experiences. It is intended that this initial framework will serve as a lens for the design, and appraisal, of sounds in interaction scenarios and that it can be iterated upon in the future through experience and empirical research.

CCS CONCEPTS

• **Human-centered computing** → *Interaction design theory, concepts and paradigms*; *HCI theory, concepts and models*; • **Applied computing → Sound and music computing**.

KEYWORDS

aesthetics, sonic interaction, sound design, user experience

ACM Reference Format:
Stuart Cunningham, Iain McGregor, Jonathan Weinel, John Darby, and Tony Stockman. 2023. Towards a Framework of Aesthetics in Sonic Interaction. In *Proceedings of Audio Mostly 2023 (AM 2023).* ACM, New York, NY, USA, 7 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Research into aesthetics has been a dominant theme in the field of Human-Computer Interaction (HCI) and user experience for over a decade, evolving and gaining increasing importance as we move through the third wave of HCI [2, 4, 9, 26, 36]. As interaction design models evolve to incorporate technological and societal changes, there is an emphasis on breaking away from traditional ways of engaging a user through language and visuals, instead focusing on experiential and emotional design qualities, oriented around perception. This is broadly described as *soma design* [25].

Although traditionally associated with the arts, the power of aesthetics in everyday objects, and by extension their interactions and related activities, should not be underestimated or trivialised. Whilst there is a philosophical and introspective nature associated with the *appreciation* of aesthetics the experience can also lead an audience into *action* [37].

The discussion within this article intends to provide a framework of sonic aesthetics. It consists of several dimensions, along with examples of their application to achieve different aesthetic results, that sonic interaction designers can consult in their activities. As such, the aim of this work is to provide an exploration of aesthetics in sonic interaction. It is oriented around devising responses to three questions:

- What are the characteristics of aesthetics in sonic interaction?
- What is as an *ugly* or *low* aesthetic in sonic interaction?
- What is a *beautiful* or *high* aesthetic in sonic interaction?

Realising answers to these questions will provide a starting point and guide for researchers and practitioners in the fields of HCI, interface design and user experience, especially those who utilise sound. This should be applicable whether sound is the sole mode of interaction or an accompaniment to other, traditional modes.

Saito [37] defined aesthetics as relating to the sensuous and design qualities (e.g., beauty, grace); characteristics held by an object of interest; or the experience of these qualities by the audience. One limitation with respect to aesthetics is that much attention has been given to *visual* aesthetics in interfaces [1, 6, 21, 32, 42] whilst other modes, especially the aesthetics of *sonic* interaction, are underrepresented. This is not to say that researchers and practitioners do not consider aesthetics in the design of sonic elements, but rather that the topic has not received the same level of formal scrutiny, possibly because their application is, as Saito [37] explains *"... being hidden in plain sight"*.

Guidelines and definitions of aesthetics, and what makes something aesthetically appealing, are rooted in traditional arts and visual design. Whilst sound and interaction design have their own sets of guidelines, principles, and heuristics, they rarely discuss aesthetics explicitly, with the exception of music [15, 20, 29, 34, 39]. It is assumed that the work of sonic interaction designers is driven

AM 2023, August 30 - September 01, 2023, Edinburgh, UK

ACM ISBN 979-8-4007-0818-3...$15.00
https://doi.org/XXXXXXX.XXXXXXX

largely by application of these guidelines, technical principles, by 'following their gut', and critical listening skills to achieve a desired aesthetic result. *In nuce*: the often-used phrase *'beauty is in the eye of the beholder'* is pertinent, though it might be better revised for our purposes *'beauty is in the* ***ear*** *of the beholder'*.

2 BACKGROUND

2.1 Aesthetics and Interaction

Hallnäs and Redström [19] strongly relate the concept of aesthetics to the notion of expression and particularly to the idea of *"expressive things"*, specifically in the context of computational objects that feature as part of everyday life. Their view, underpinned by the association that strong aesthetic design is a matter of logic, recognises a dual, interconnected, purpose of interface components: that they serve as an interaction *mechanism* as well as a tool of *expression* and that each complements the other by their association. Whilst generally oriented to the notion of computing *devices* and physical *"things"* that inhabit a human-computer ecosystem where technology is ubiquitous, the role of the interface is an important feature of such objects. This is partly identified by Hallnäs and Redström, identifying that it becomes necessary to consider how usability may need to be re-framed when we consider the *meaningfulness* of a technology in our lives. There is something of a boundary drawn between objects that are *tools* and those that are part of a person's *life*. The authors suggest that only in the case of the latter would aesthetics deliver greatest benefit. Although it may be the case that by delivering an aesthetically pleasing experience, something that began as a *tool* may transcend to become a *meaningful* element.

In describing the core aspects of emotional design, Norman [33] notably identifies three key components: *visceral*; *behavioural*; and *reflective*. Each incorporates aspects of enjoyment and pleasure that a person derives from their experience with an artifact. Norman's work is especially useful in that it incorporates consideration of whether *form* (visceral) or *function* (behavioural) is the greater contributor to the aesthetic experience of an object, whilst also providing space for exploration and study of the object and its *interaction* (reflective).

Lavie and Tractinsky [27] deal with the possible tension that can exist for interaction designers in their work as they follow established rules and conventions, whilst at the same time are attempting to demonstrate novelty. Both are considered as important in achieving high aesthetic affect, but naturally they may be at odds with one another, since breaking rules is often considered necessary in order to achieve originality. In doing so, the authors consider studies of web sites, and their varying degrees of visual design and intended aesthetic qualities in order to develop scales for measuring users' aesthetic experiences. Their work considers the established design rules, termed *classical aesthetics*, whilst the breaking of such rules to demonstrate originality is labelled *expressive aesthetics*. Despite being oriented toward visual design, such a distinction is useful, since it allows us to consider that classical aesthetics may provide a route to create aesthetically pleasing outcomes, likely to have broad appeal, whilst expressive aesthetics may produce greater subjective appraisal and be more polarising. Correspondingly, their scales realised for measuring the factor of classical aesthetics contain items such as: *pleasant design*, *clean design*, and *symmetric design*. Items in the factor of expressive aesthetics included: *fascinating design*, *creative design*, and *original design*. These scales may be broadly transferable into other modes, such as sound, although modification and revaluation would be necessary first.

The Unified Model of Aesthetics (UMA) [22] recognises the tension that exists regarding the concept of novelty and its influence on an object (or interaction) being perceived as pleasing or displeasing. The UMA operates using a set of three dimensions, with each extreme embodying an overarching concept relating to a human goal or intention. In brief, the goals of a person relating to *safety* are characterised by ends of each dimension labelled connectedness, typicality, and unit. Whilst *accomplishment* is related to the opposite ends of the three the dimensions labelled autonomy, novelty, and variety. The model has subsequently been evaluated in more detail, suggesting that it may be valid [7].

Notably, the UMA suggests that a fine balance must be struck between the experience of something being either typical or novel and that preference of an individual is likely to be influenced by their current situation or context. This situational factor is likened to the concept of safety or risk aversiveness. In other words, people prefer typical or familiar aesthetic experiences in high-risk settings, whilst they are more welcoming of new or unfamiliar experiences when any risk is minimal, or they feel comfortable.

Tractinsky *et al.*'s [43] work *"what is beautiful is usable"*, is considered an important milestone in research that relates aesthetics to usability. In their study, the authors evaluate three different levels of aesthetic design, previously characterised by the aesthetic appraisal of interfaces by study participants, which contained identical components and features, but differing layouts and locations of objects in two-dimensional space. Their work showed that perceptions of aesthetic quality and usability were related both pre and post use of the system trialled in the experiment. In discussing future developments and factors to consider, the authors identify that a key question in this field is that of *"What makes an aesthetic interface?"*.

In a later study that sought to examine the relationship between interface aesthetics and interface usability, Tuch *et al.* [44], performed an experiment based upon participants performing tasks in an online shopping scenario. One of the independent variables was aesthetics (low and high), with the intention of being able to present users with a *"beautiful"* and *"ugly"* version of the web site. Selection was drawn from a set of templates, based upon the judgement of a panel of experts. So as not to accidentally interfere with the designed usability, the authors manipulated only the *background colour* and *texture* of the web sites, along with decorative *images*, to facilitate two aesthetic versions. These features can be considered low-level or basic characteristics of visual interface design. Furthermore, the authors highlight examples of other research studies where, in manipulating levels of visual interface aesthetics, authors have tended to alter features such as the *location* of buttons, *visual design*, *skins*, *colour*, *layout*, and *font*. These are intuitive and well-defined characteristics, at the heart of visual interface design choices.

As with similar studies, Tuch and colleagues [44] measured participants' perceived usability of each version of a software interface as well as its perceived aesthetic quality. The latter can be particularly important in determining if the different aesthetic levels, devised by the researchers, is recognized by participants, and may be considered further evidence that interface aesthetics are a subjective phenomenon – which might be broadly agreed upon within some limit, but is not necessarily universal. A salient point made in this work, is that the relationship between aesthetics and usability can be intertwined. For instance, consider the staggered placement of a series of buttons in 2D space, as opposed to them being presented in a regimented set of tiles. The former may be considered a poor aesthetic choice, due to the breaking of symmetry but also poor usability, since the user must perform more fine adjustment of their pointing device, as described by Fitts' Law [14].

2.2 Aesthetics and Sonic Interaction

Considering the production of earcons [8], which are structured, abstract sounds, often with musical components, a set of experimentally derived guidelines were produced by Brewster *et al.* [10]. These guidelines provided a set of parameters that an interaction designer can manipulate, namely: *timbre*; *register*; *pitch*; *rhythm*; *intensity*; and *spatial location*. It is noted that high intensity and high pitch are common and direct ways of being able to grab attention of a user, whilst *combinations* of other parameters may lead to similarly attention-grabbing outcomes. In the real world, high-pitch sounds are often correlated to danger or potential damage, used to attract the attention of a person [13, 35], and may improve task focus [5]. However, high levels of attention-grabbing sound may not always an optimal route to take depending upon the context. For example, in hospital intensive care units, such alerts may disturb patients and staff or have detrimental effects on the mental well-being of patients [12].

These guidelines provide a useful summary of features available to the sonic interaction designer, although any indication as to how they might be combined to achieve aesthetically pleasing results is not explicit. However, the broad notion of aesthetics may be inferred as being relevant by the work of Blattner *et al.*, [8] who explained that an earcon *"…should not be unpleasant nor fatiguing"*.

While Brewster *et al.* [10] may not discuss how sounds may be combined to provide high or low aesthetics, elsewhere in the field of electroacoustic composition, Smalley's [40] concept of *spectromorphology* discusses the way in which spectral characteristics of sounds are manifested in time. Smalley's discussion may be informative when we wish to think about how different components of sound can be combined, since he allows us to breakdown sounds into onset, continuant and termination phases. Here Smalley is not interested in high/low aesthetics in the terms we have described here, yet it is possible that the principles outlined by Brewster *et al.* [10] might fruitfully be extended through consideration of the spectromorphological properties of earcons. Later work by Hoggan and Brewster [24], which draws upon related discussions of Buxton [11], identifies that some sounds in interfaces can annoy users, essentially providing a poor experience and, we argue, directly related to a poorer or lower quality aesthetic. *Noise* is cited as being problematic, at least in the sense of a signal-to-noise ratio, where the noise is any distraction that masks the intended meaning of a sound. Such noise might be related to the situation or context in which a user finds themselves, or could be part of the sound itself, either intentionally or unintentionally. Other examples discussed by Hoggan and Brewster [24] relate to sounds being too *loud*.

Mynatt [31] provided similar recommendations in the creation of auditory icons [16, 17], which are typically sampled sounds that represent real-world objects or events, being skeuomorphic in nature and may give rise to positive emotions, perhaps relating to sensations of nostalgia [3]. These guidelines utilised broad design principles and practices, being an iterative process involving evaluation, rather than specific parameters that the interaction designer may control. However, removing the evaluation elements, the sound design guidelines can be related to the sound characteristics of: *duration*; *bandwidth*; *intensity*; and *quality*.

2.3 Aesthetics and Visual Design

In attempting to achieve varying degrees of visual aesthetic, it appears fortuitous to apply frameworks or guidelines for their design, such as *Gestalt principles*, which are encountered in the design of visual interfaces for interactive systems [18].

Studies in the domain of product design, found that factors of *unity* and *prototypicality* strongly related to aesthetic perceptions [45]. Although neither is defined in classical Gestalt theory, the authors recognised that *unity*, the grouping or joining of visual elements in a design, draws upon multiple Gestalt theories of proximity, similarity, and common fate, and that these are contributory. Conversely, *prototypicality* has no clear direct relation to Gestalt principles, but rather is the extent to which a design best represents the category or nature of object (or phenomenon) that has been designed. To this end, one might consider this as being an indication of how well a new design meets the *expectation* of its audience or user.

Lim *et al.* [28] took this concept further and considered how Gestalt principles may shape all elements of *interaction*, not only those concerned with visual *interface* design. Their work specifically considered the back-and-forth nature of interaction with an artifact, where feedback is provided over a series of user actions. Although their work was not specifically directed towards sonic interaction, they described the concept of interaction Gestalt as being composed of three attributes of *shape*, *describe*, and *analyze* and distinguished this notion of being separate, but related to, the properties of the artifact itself and that of the user's perceptions and *experiences* of the artifact.

3 SOUND AND AESTHETICS

3.1 Design

Saito [37] cautions against the application of traditional art standards in the examination of aesthetics in other domains, specifically everyday objects, by applying a *"mono-framework for aesthetic discourse"*. They explain that is due to mismatches that occur due to the differences in purpose, context, and a shift in balance between contemplation and action. This is salient advice in attempting to define aesthetic characteristics in sound for the purposes of interaction, although it does not exclude using traditional art and visual rules as a baseline from which sonic aesthetics may be described.

In designing interactions, Lim *et al.* [28] provide a useful reminder that, to achieve some aesthetic effect, we must be aware of what we have the potential to manipulate (as well as requiring the skill to achieve the desired effect). In the case of sonic interaction, this can be taken as referring to the acoustic parameters and operations available to the practiced sound designer. The work of Tuch *et al.* [44] showed that simple and fundamental characteristics of visual interface elements can reflect different levels of aesthetic. By extension, sonic interface designers might consider features such as *loudness*, *envelope*, *panning*, *pitch*, and *timbre* as broadly equivalent.

Scott-James [38] shares this notion that there are parallels between the design elements of images and sound, stating *"Image elements have similar sound equivalents; color and hue are akin to frequency, pitch, or timbre"*. This is in addition to explaining that much of the work of designing sonic aesthetics, at least in the context of film, is about being able to realise an underpinning, intended *"feel"* or emotion. This leads us to argue that sounds designed to be pleasant or stimulate a positive emotion (such as happiness or excitement) are likely to be considered as having a high aesthetic, whilst those designed to provoke negative emotions (fear or disgust, for instance) may be perceived as being of a low aesthetic.

Soundscapes and their sonic components can generally be perceived as aesthetically pleasing, displeasing or even neutral [41]. Whether a hedonic or aversive experience is elicited is predominantly based on prior *experiences* and *associations* [30]. However, when sounds are combined pleasing sounds can become displeasing, as can displeasing sounds become pleasing. Once a meaning has been attached to a sound, a previously pleasing sound might change to a displeasing one, and vice versa.

3.2 Principles

Zwicker and Fastl [46] recognised that features of a perceived sound relate to its perception of pleasantness. In doing so, they produced a formulation for calculating the perceived sensory pleasantness, based upon factors of sharpness, roughness, tonality, and loudness. It is useful to describe these characteristics in further detail and the underlying phenomenon.

There are common underlying principles that affect how pleasing or displeasing a sound might be. The first is how *loud* a sound is, generally louder is more displeasing. Quiet auditory environments that are also considered pleasing are thought to promote proactive behaviour. If sounds are considered relevant in a particular context and they support what the other senses are experiencing without dominating, then the sounds are normally described as pleasing. In contrast, a loud sound is often considered a form of interference or noise and is almost uniformly accepted to have an inherent impact upon quality of life, and therefore has an association with being displeasing. A sound does not have to be loud to be displeasing, a quiet sound when unwanted in context can be just as annoying. The issue is related to interruptions of mind or distraction, which can lead to attentional fatigue. *Pleasant* sounds are often associated with a safe environment, which facilitates relaxation with a positive effect on long-term health. Unpleasant sounds can be related to danger, and require constant reactive attention, which creates stress and if there is prolonged exposure can have a negative effect on long-term health.

If a sound contains an excess of acoustic energy between 2-4 kHz it is often described as being 'sharper'. This sharpness is the main reason why some sounds are considered annoying and unpleasant. Screeching birds, squealing brakes and the often-cited fingernails down a blackboard all contain high levels of 2-4 kHz. The human hearing system is most sensitive to 2-4 kHz, and high levels of these frequencies can potentially damage the relevant hair cells in the cochlea much more easily than in any other *frequency* range. By making 2-4 kHz less prominent the perceived unpleasantness can be reduced, suggesting that the aversion is closely linked to a natural unconscious desire to protect hearing apparatus and avoid pain or discomfort.

If a sound fluctuates in *intensity* between the rates of 20 to 200 times per second, it can produce a roughness that affects the frequency balance or timbre of a sound. Some car manufacturers will actually test each car individually on a specialist test track as it comes off the production line to check that no parts are loose and rattling, even at very quiet levels, so that the new owner does not experience any dissatisfaction with their new vehicle. Sounds that are created by slipping and gripping or scraping often have a fluctuation in volume that is perceived as roughness, and therefore as unpleasant. A knife on a bottle has been considered the most unpleasant sound, closely followed by a fork on glass and ruler on a bottle, all of which generate sound through scraping actions. Slower sounds with predominantly low pitches are often thought to be pleasing, especially when associated with natural world, such as wind, waves and rain.

There is a general preference for *clarity of pitch*, and this can be achieved through averaging. By recording multiple versions of a voice in time with itself and layering them imperfections can be reduced. In a similar manner adding reverberation will make the harmonics louder and the noise components quieter, resulting in a smoother pitch. The complex reflections associated with reverberation apply short delays, spatial width and height, all of which when combined with the different absorption and reflection coefficients combine to create what is often referred to as a fuller, richer sound. The closer to the average that a voice falls for each gender the more aesthetically appealing it is often perceived to be, especially in terms of its formants.

Consonance can be defined as musical notes that are harmonious when played together. Consonance is not confined to music; it can be experienced in any combination of sounds that are spectrally either far apart or identical. Dissonant sounds are those that are spectrally very similar but not identical, producing a beating effect due to phase shifts, that are sometimes heard as an unpleasant roughness. Another effect is being heard in that dissonant sounds have partials that are not integers of the fundamental frequency or first partial, which are called inharmonics. In contrast consonant sounds have partials that are integers of the fundamental frequency, and these are termed harmonics. Few sounds are purely harmonic or inharmonic, often they fall somewhere in between, and by varying the amount of integer partials (harmonics) and non-integer partials (inharmonics) the perceived aesthetics of a sound or combination of sounds can be affected.

4 A FRAMEWORK OF SONIC AESTHETICS

Using the literature and discussion so far, in Table 1, we propose key dimensions of sound, synthesised from the literature and our own experiences in the fields of sound design and HCI research. Each dimension is described, along with typical configurations, that we believe can result in the perception of a low or high aesthetic experience for an audience or user. Of course, the perception of aesthetic quality is unlikely to be manifest itself in such a dichotomous manner in the real-world, but rather in a wide-ranging or continuous fashion.

As such, our intention is to demonstrate what the extremes might be like, as opposed to providing an exhaustive definition. It should also be noted that the use of the spatial dimension may not always be relevant or possible, depending upon the situation of use and reproduction equipment available. Since sound is a time-based medium, it is expected that the qualities within each of these dimensions may purposefully dynamically alter over the duration of a sound, if this is the intention of the interaction designer and that this may account for Smalley's [40] *spectromorphology*.

To help exemplify the use of the dimensions in the framework, the low and high aesthetic columns of Table 1 provide a short description of a sonic interaction scenario (in parentheses). These illustrate how the aesthetic extremes of each dimension may manifest themselves in a scenario. In this case, we ask the reader to consider the sound that the buttons of an Automated Teller Machine (ATM) may make whilst the user is entering their PIN (Personal Identification Number) prior to performing a transaction.

We suggest that the dimensions of *intensity*, *pitch*, *timbre*, *spatial*, *fidelity*, and *context* are closely related to Lavie and Tractinsky's [27] classical aesthetics, whilst *expectation and originality* deliver scope to achieve expressive aesthetics. The latter dimension is especially challenging to define, perhaps indicative of its somewhat nebulous and highly subjective manner. This dimension consciously includes the concepts of originality and expectation, since the two seem largely inseparable in this domain. Part of the essence of rule breaking is about defying what is expected. Such rules may quickly result in listener fatigue and sounds being perceived as clichéd. For instance, tinnitus sounds (constant ringing tones) in video games to convey poor health or luxurious reverb effects added to voices in a dream sequence. The auditory elements are uninspired but pertinent and fit the stereotype.

Regarding the notion of a user's expectation of sonic interaction, this may relate to a single *instance* (e.g., a sound being played as feedback to a key press) or to the overall experience of a particular *journey*, or *series of journeys* (e.g., all sounds heard during a series of interactions that take place during a task). Particularly when dealing with a system's overall scheme, we suggest that this is viewed as its *potential prototypicality*: the potential of the sound(s) to meet the user's expectations of a given interactive system.

To meet these expectations requires a coherence within the overall scheme, which the user can learn from (gradually, without effort), and that can be drawn upon when they encounter new scenarios. Achieving such an aim leads us towards sonic interaction in a system achieving a form of *unity*. Unity is unlikely to be achieved on a sound-by-sound basis or an instantaneous quality of a particular scheme or system of sounds and one which is learnt via a meaningful structure embedded within the sonic interaction design.

5 DISCUSSION AND FUTURE WORK

The framework presented in Table 1 serves to define and illustrate our current view of sonic interaction aesthetics. It provides a direct response to each of the three questions posed in the Introduction of this article. We intend that the framework is the *beginning* of a process and point for discussion and refinement, rather than a definitive or conclusive *outcome*. However, it should serve as a lens, and set of guidelines, through which sonic aesthetics might be both designed and appraised. It draws upon established principles and recommendations from the technicalities and practicalities of sound design, whilst considering the user perception and experience of its dimensions.

It may be reasonable to suggest that a great deal can be done in sonic interaction aesthetics solely making use of *classical aesthetics* and dimensions of the framework that manifest themselves as low-level features and parameters of sound (e.g., intensity, pitch, timbre). This is especially so when taking the position that the intention of sonic interaction is rarely, if ever, purely for the sake of art alone, but to provide a usable and productive outcome for the user. However, this does not mean that *expressive aesthetic* qualities should be neglected or ignored, particularly if the dimensions of expectation, providing unity, can be aligned with originality. For instance, an experience of a sonic interaction scheme that consistently and predictably breaks the rules may be judged as highly aesthetically pleasing, whilst also providing the required usability and task performance.

One way to resolve the discrepancy between the dimensions of the proposed framework, which sit within classical aesthetics, and those from expressive aesthetics, may come via an analysis of the sounds themselves. For instance, sounds designed to sit within a scheme could be subjected to a cluster analysis [23], based on the dimensions of the framework, to determine the extent to which they achieve the desired sound-by-sound aesthetic intent, whilst providing an insight as to the degree to which they conform to an *expectation* of unity. Analysis of this kind would also allow the sonic interaction designer to assess which sounds were outliers, potentially providing an indication of their rule-breaking characteristics, which have potential for *originality*.

Future activities in this field should consider eliciting aesthetic judgements of sounds in interaction scenarios that exemplify a broad spectrum of material according to the framework presented, as others have done with visual user interface designs in the literature. Obtaining such empirical data would provide credence to the framework and help determine if its dimensions are valid predictors of aesthetic appeal. A natural expansion of this would be to consider sonic elements on both the sound-by-sound as well as the overall scheme perspectives, as outlined earlier.

Consideration should also be given to the spectromorphology of a sound or scheme of sounds, and how sounds evolve and change over their duration. This has not been explicitly considered in our discussion so far, beyond the acknowledgement that sounds exist in the temporal domain for the user. Many natural sounds that draw upon the principle of sonic icons for instance may contain

Table 1: A Framework of Sonic Aesthetics

Dimension	Description	Low Aesthetic (example with ATM interaction)	High Aesthetic(example with ATM interaction)
Intensity	The perception of loudness of the sound.	Overly loud. Inconsistent. (Envelope alters significantly, at its loudest is uncomfortable, drawing attention to itself).	Quiet but clear enough to be perceived as intended. Consistent. (Envelope of the sound is minimal and is comfortably perceived).
Pitch	The fundamental frequency, or musical note, dominant in the sound.	Pitches at the edges of being perceived by the human auditory system. Excess use of pitches in areas of high sensitivity. Rapid changes over time. (A low pitch tone is played (100 Hz) that rapidly descends).	Pitches comfortably in the human auditory system range and changes that are paced appropriate to the sound duration. (A static pitch is applied in fourth of fifth octave (250 Hz – 1000 Hz)).
Timbre	The character of the sound and complexity of its frequency spectrum.	Sounds are dissonant and /or use a narrow frequency spectrum. (A sound with a small number of harmonics, random, non-integer, multiples of the fundamental frequency).	Sounds are consonant and/or use a broad frequency spectrum. (A harmonically rich sound, applying integer multiples of the fundamental frequency, such as a square or pulse tone).
Spatial	The position of the sound relative to the listener's position.	Playback is located at an extremity, in the case of low channel numbers. Playback comes from a location unrelated to the position of the artefact with which they are interacting, in the case of large numbers of channels and/or three-dimensional capability. Dynamic panning is erratic. (Sound playback comes from beyond the edges of the ATM machine and moves across the azimuth of the user's perception).	Playback is subtly panned in the available number of channels, relative to the position of the artifact that is being interacted with. Use of dynamic panning is smooth and slow, relative to the sound duration. (The perceived location of playback comes from the ATM machine itself throughout the interaction).
Fidelity	The production quality of the sound. Considers the clarity of the sound and its intention, including the presence of noise and other artefacts and the discernability between signal and noise.	The sound contains excessive amounts of noise or distortion and may be glitchy in its playback. (Audible, broad-spectrum noise or hiss is present that masks the intended pitch of the tone).	The sound is clear and absent of distracting artefacts. (The intended sound's pitch is clearly audible along with other content. Dynamics are perceived easily. Intended noise, such as buzzing is in limited frequencies).
Context	The relationship between the sound, interaction task, and any other sensory stimulations (e.g., touch, visual, olfaction). Context may not always be known to the designer (such as for mobile applications).	The sound associated with the action is perceived as inappropriate or unsuitable to the user. The sound violates the expectation of the user and may conflict with the other stimuli of the task. (A random number of short duck quack sounds are emitted when the user presses a key).	The sound meets the expectations of the user, being relatable to the current task and perceived to fit it well. The sound used is familiar and organic. (A plain but clear bleep sound is utilised, which behaves consistently between presses and like previous interactions. Interaction with intensity is important, insomuch as the sound must be loud enough to be clearly perceived over any background noise.)
Originality and Expectation	The level of predictability and familiarity in terms of the overall context of use, as well as in general experience.	Clichéd, overused sound. (Electronic, synthesised, monotonic 'beeps').	Stimulates engagement through novelty whilst retaining relevance. (A suitably randomised set of cues that convey the full context of the actions, such as virtual keyboard keystrokes varying when entering numbers, cash register 'pings' when choosing to withdraw currency, a piggy bank noise when making a deposit, etc.).

their own inherent envelope and spectromorphology. However, if sampled sounds are to be combined and/or synthesised sounds or earcons employed, then there is increased scope for this to be an additional dimension of the framework.

REFERENCES

[1] Ahamed Altaboli and Yingzi Lin. 2011. Investigating effects of screen layout elements on interface and screen design aesthetics. *Advances in Human-Computer Interaction* 2011 (2011), 1–10.

[2] Tomás Alves, Joana Natálio, Joana Henriques-Calado, and Sandra Gama. 2020. Incorporating personality in user interface design: A review. *Personality and Individual Differences* 155 (2020), 109709.

[3] Michael L Austin. 2016. Safe and sound: Using audio to communicate comfort, safety, and familiarity in digital media. In *Emotions, technology, and design*. Elsevier, 19–35.

[4] 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*. Association for Computing Machinery, New York, NY, United States, 2689–2698.

[5] Anil Ufuk Batmaz and Wolfgang Stuerzlinger. 2021. The effect of pitch in auditory error feedback for fitts' tasks in virtual reality training systems. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. IEEE, IEEE, 85–94.

[6] Michael Bauerly and Yili Liu. 2008. Effects of symmetry and number of compositional elements on interface and design aesthetics. *Intl. Journal of Human–Computer Interaction* 24, 3 (2008), 275–287.

[7] Michaël Berghman and Paul Hekkert. 2017. Towards a unified model of aesthetic pleasure in design. *New Ideas in Psychology* 47 (2017), 136–144. https://doi.org/10.1016/j.newideapsych.2017.03.004

[8] Meera M Blattner, Denise A Sumikawa, and Robert M Greenberg. 1989. Earcons and icons: Their structure and common design principles. *Human–Computer Interaction* 4, 1 (1989), 11–44.

[9] Letizia Bollini. 2017. Beautiful interfaces. From user experience to user interface design. *The Design Journal* 20, sup1 (2017), S89–S101.

[10] Stephen A Brewster, Peter C Wright, and Alistair DN Edwards. 1995. Experimentally derived guidelines for the creation of earcons. In *Adjunct Proceedings of HCI*, Vol. 95. Cambridge University Press, 155–159.

[11] William Buxton. 1989. Introduction to this special issue on nonspeech audio. *Human Computer Interaction* 4 (1989), 1–9.

[12] Julie L. Darbyshire, Paul R. Greig, Lisa Hinton, and J. Duncan Young. 2021. Monitoring sound levels in the intensive care unit: A mixed-methods system development project to optimize design features for a new electronic interface in the healthcare environment. *International Journal of Medical Informatics* 153 (2021), 104538. https://doi.org/10.1016/j.ijmedinf.2021.104538

[13] Judy Edworthy and Neville Stanton. 1995. A user-centred approach to the design and evaluation of auditory warning signals: 1. Methodology. *Ergonomics* 38, 11 (1995), 2262–2280.

[14] Paul M Fitts. 1954. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology* 47, 6 (1954), 381.

[15] Guy E Garnett. 2001. The aesthetics of interactive computer music. *Computer Music Journal* 25, 1 (2001), 21–33.

[16] William W Gaver. 1989. The SonicFinder: An interface that uses auditory icons. *Human–Computer Interaction* 4, 1 (1989), 67–94.

[17] William W Gaver. 1993. Synthesizing auditory icons. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*. Association for Computing Machinery, New York, NY, United States, 228–235.

[18] Lisa Graham. 2008. Gestalt theory in interactive media design. *Journal of Humanities & Social Sciences* 2, 1 (2008).

[19] Lars Hallnäs and Johan Redström. 2002. From use to presence: on the expressions and aesthetics of everyday computational things. *ACM Transactions on Computer-Human Interaction (TOCHI)* 9, 2 (2002), 106–124.

[20] Eduard Hanslick. 1891. *The beautiful in music: A contribution to the revisal of musical aesthetics*. Novello.

[21] Jan Hartmann, Alistair Sutcliffe, and Antonella De Angeli. 2008. Towards a theory of user judgment of aesthetics and user interface quality. *ACM Transactions on Computer-Human Interaction (TOCHI)* 15, 4 (2008), 1–30.

[22] Paul Hekkert. 2014. *Aesthetic responses todesign: a battle of impulses*. Cambridge University Press, 277–299. https://doi.org/10.1017/CBO9781139207058.015

[23] Christian Hennig, Marina Meila, Fionn Murtagh, and Roberto Rocci. 2015. *Handbook of cluster analysis*. CRC press, Boca Raton, FL, United States.

[24] Eve Hoggan and Stephen Brewster. 2012. Nonspeech auditory and crossmodal output. In *Human Computer Interaction Handbook : Fundamentals, Evolving Technologies, and Emerging Applications* (3 ed.), Julie A. Jacko (Ed.). CRC press, Boca Raton, FL, United States, Chapter 10, 211–235.

[25] Kristina Hook. 2018. *Designing with the body: Somaesthetic interaction design*. MIT Press, Cambridge, MA, United States.

[26] Kasper Hornbæk and Morten Hertzum. 2017. Technology acceptance and user experience: A review of the experiential component in HCI. *ACM Transactions on Computer-Human Interaction (TOCHI)* 24, 5 (2017), 1–30.

[27] Talia Lavie and Noam Tractinsky. 2004. Assessing dimensions of perceived visual aesthetics of web sites. *International journal of human-computer studies* 60, 3 (2004), 269–298.

[28] Youn-kyung Lim, Erik Stolterman, Heekyoung Jung, and Justin Donaldson. 2007. Interaction gestalt and the design of aesthetic interactions. In *Proceedings of the 2007 conference on Designing pleasurable products and interfaces*. Association for Computing Machinery, New York, NY, United States, 239–254.

[29] Edward A Lippman. 1994. *A history of Western musical aesthetics*. U of Nebraska Press.

[30] Josh H McDermott. 2012. Auditory preferences and aesthetics: Music, voices, and everyday sounds. In *Neuroscience of preference and choice*. Elsevier, 227–256.

[31] Elizabeth D Mynatt. 1994. Designing with auditory icons: how well do we identify auditory cues?. In *Conference Companion on human factors in computing systems*. Association for Computing Machinery, New York, NY, United States, 269–270.

[32] David Chek Ling Ngo, Lian Seng Teo, and John G Byrne. 2003. Modelling interface aesthetics. *Information Sciences* 152 (2003), 25–46.

[33] Donald A Norman. 2005. *Emotional design: Why we love (or hate) everyday things*. Basic Books.

[34] Max Paddison. 1997. *Adorno's aesthetics of music*. Cambridge University Press.

[35] Roy D Patterson. 1990. Auditory warning sounds in the work environment. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences* 327, 1241 (1990), 485–492.

[36] Ingrid Pettersson, Florian Lachner, Anna-Katharina Frison, Andreas Riener, and Andreas Butz. 2018. A bermuda triangle? A review of method application and triangulation in user experience evaluation. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. Association for Computing Machinery, New York, NY, United States, 1–16.

[37] Yuriko Saito. 2007. *Everyday aesthetics*. Oxford University Press on Demand.

[38] Kahra Scott-James. 2018. *Sound Design for Moving Image: From Concept to Realization*. Bloomsbury Publishing.

[39] Roger Scruton. 1997. *The aesthetics of music*. Oxford University Press.

[40] Denis Smalley. 1986. *Spectro-morphology and Structuring Processes*. Palgrave Macmillan UK, London, 61–93. https://doi.org/10.1007/978-1-349-18492-7_5

[41] Xiuhua Song, Xinbo Lv, Dongming Yu, and Qianqian Wu. 2018. Spatial-temporal change analysis of plant soundscapes and their design methods. *Urban Forestry & Urban Greening* 29 (2018), 96–105.

[42] Noam Tractinsky. 1997. Aesthetics and apparent usability: empirically assessing cultural and methodological issues. In *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems*. Association for Computing Machinery, New York, NY, United States, 115–122.

[43] Noam Tractinsky, Adi S Katz, and Dror Ikar. 2000. What is beautiful is usable. *Interacting with computers* 13, 2 (2000), 127–145.

[44] Alexandre N Tuch, Sandra P Roth, Kasper Hornbæk, Klaus Opwis, and Javier A Bargas-Avila. 2012. Is beautiful really usable? Toward understanding the relation between usability, aesthetics, and affect in HCI. *Computers in Human Behavior* 28, 5 (2012), 1596–1607.

[45] Robert W Veryzer Jr and J Wesley Hutchinson. 1998. The influence of unity and prototypicality on aesthetic responses to new product designs. *Journal of consumer research* 24, 4 (1998), 374–394.

[46] Eberhard Zwicker and Hugo Fastl. 1999. *Psychoacoustics Facts and Models Second Updated Edition*. Psychoacoustics, Vol. 22. Springer Berlin / Heidelberg, Germany.