Sonification of Emotion I: Film Music

Emery Schubert[1] *Sam Ferguson*[1] *Natasha Farrar*[1] *Gary E. McPherson*[2]

[1]Empirical Musicology Lab,
School of English, Media and Performing Arts,
University of New South Wales, Sydney, 2052,
Australia
e.schubert@unsw.edu.au
samuel.ferguson@unsw.edu.au

[2]Melbourne Conservatorium of Music,
University of Melbourne,
Victoria, 3010, Australia
n.farrar@student.unsw.edu.au
g.mcpherson@unimelb.edu.au

ABSTRACT

This paper discusses the uses of sound to provide information about emotion. The review of the literature suggests that music is able to communicate and express a wide variety of emotions. The novel aspect of the present study is a reconceptualisation of this literature by considering music as having the capacity to sonify emotions. A study was conducted in which excerpts of non-vocal film music were selected to sonify six putative emotions. Participants were then invited to identify which emotions each excerpt sonified. The results demonstrate a good specificity of emotion sonification, with errors attributable to selection of emotions close in meaning to the target (excited confused with happy, but not with sad, for example). While ‘sonification’ of emotions has been applied in opera and film for some time, the present study allows a new way of conceptualizing the ability of sound to communicate affect through music. Philosophical and psychological implications are considered.

1. INTRODUCTION

According to Kramer et al. [1] “sonification is the transformation of data relations into perceived relations in an acoustic signal for the purposes of facilitating communication or interpretation” (p. 4). Of the more practically oriented applications, the display of numerical data has been obvious and frequent, from Geiger counter inspired techniques [2] to sonification of stock exchange reports [3]. However, this paper discusses the application of auditory display of something that is supposedly difficult to quantify – emotions (for a review, see [4]). Our approach is inspired by Vickers and Hogg [5] who remark that “through the intervention of the musician, the [musical instrument] renders in sound (albeit in a highly complex and abstract way) the score, the technique of the musician, the physics of the [instrument], the emotional state of the musician and the composer” (p. 211).

Hermann’s definition of sonification [4] defines four essential attributes of a sonification: 1) it reflects objective properties or relations in the input data, it has both a 2) systematic and 3) reproducible transformation to sound, and 4) it can be used with different input data. This definition delineates the difference between sonifications that focus on representing input data for analysis and description, and other audio works that may represent data for artistic or aesthetic purposes, but without a strong focus on analysis of the data.

The case will be argued here that music can be considered a culturally calibrated medium for the signification of some emotions, and that referring to music as an auditory display of emotion could provide some interesting alternatives to philosophical debates about emotional expression in music and to psychological paradigms.

This will be followed by the reporting of an experiment where an individual is given the task of describing the emotions that are sonified by film music excerpts. Furthermore, the use of emotional content of music as a mapping target is discussed, and possible speculative techniques for the encoding of arbitrary information, (rather than only music’s emotional content) will be outlined.

2. PHILOSOPHY AND PSYCHOLOGY OVERVIEW

Sonification of emotion by music actually has a long history, and is usually referred to as ‘emotional expression’ in music. Music is believed to express or communicate emotions, such as joy, excitement and sadness to the listener. For example, Hevner [6] developed a circle of adjectives that were organized into groups of words (clusters) of similar emotional meaning (e.g. merry and joyous in one cluster, versus somber and sober in another). A piece of music could then be classified as expressing one or more emotion terms. A further refinement of this approach was to have each adjective rated for its appropriateness [7], and to have them rated continuously as the music unfolded [8]. A common thread in this research is the attempt to identify how musical structure encodes the emotion, which is then decoded by the listener. We will therefore refer to this as the *decoding paradigm* of emotion in music research. An important developer of decoding theory in recent times has been Juslin who proposed the Lens model and the GERMS model [9] to explain how the emotion gets ‘into’ the musical structure, to then be decoded by the listener. About the Lens model, Juslin writes:

> This model is meant to illustrate how performers encode (i.e., express) emotions by means of a

> number of probabilistic (i.e., uncertain) but partly redundant cues (i.e., sources of information). The emotions are decoded (i.e., recognized) by listeners who use these same cues to judge the emotional expression. The cues are probabilistic in the sense that they are not perfectly reliable indicators of the intended emotional expression. [10, p. 1798]

One way the Lens model, and more generally the decoding paradigm has been investigated experimentally is for the performer to be instructed to imbue a piece with a particular emotion, and the listener then tries to decode the expressed emotion by rating a number of emotions. Typically if the 'correct' or intended emotion is not selected or rated as the emotion that best describes the music, it is referred to as inaccurate or a confusion [see for example 10-14]. While not a necessary limitation of the decoding paradigm, the focus on finding a particular musical structure that encodes a putative emotion may be, in part, responsible for some limitations in how emotion in music is treated in the laboratory. That is, much of this kind of research limits what is considered a correctly decoded emotion to enable a focus on a set of musical features and characteristics that convey that target emotion. For example, sad music is slow, soft and in the minor mode [for a literature review, see 15].

However, sonification was a term that would not have sat well with those who believed that music was unable to express anything directly, and at the same time could express anything at all. The mapping of emotion to music, according to this view, would be arbitrary and in some cases meaningless. In fact, Langer argues that this lack of specificity is one of the advantages that music has over language. "What is here criticized as a weakness, is really the strength of musical expressiveness: that *music articulates forms which language cannot set forth"* [16, p. 233]. Langer's views have been influential in music philosophy [e.g. 17], despite research in music psychology having quantified emotion based on psychological understandings of the structure of emotion.

These two epistomologies - the rationalist philosophical and the reductionist psychological - can to some extent be reconciled by thinking of music as having the ability to sonify emotion. First, if we accept that there are cultural factors that constrain the possibly chaotic aspects of musical associations [18], we immediately limit the number of meanings that music is likely to express to any individual, whether ineffable or reportable. Further, the wide ranging disagreements among philosophers could, to some extend be reconciled by variables such as individual differences [19-21]. Second, the reductionism of the psychological approaches, in particular the approaches using the *decoding paradigm,* may be seen as an approximation of the true, complex, and at times ineffable experience. For example, if a participant reports that a piece of music expresses a score of 7/10 for sadness, it does not literally mean that the individual has the equivalent, but reduced, experience, but rather that given the limited choice, the best answer the listener could give to reflect the complex experience was 7/10 for sadness. Further, if more than one emotion is rated as being expressed by the music, it may be possible that both emotions have in some way or combination been projected, rather than one or more being incorrect or confusions [22].

This is where we argue that understanding music as the sonification of emotion helps to bridge the discrepancy between emotional expression in music that appears in psychological versus philosophical writings. By asserting that music is (among other things) an auditory display of emotions, we should be able to measure how well music is able to perform this sonification experimentally. It may be seen as an alternative to the decoding paradigm, or as a refinement. By drawing on the typical applications of auditory display, we posit that a sonified emotion through the auditory medium, music, is a representation of the emotion expressed by the music. When assessing the success of the sonification the reduction and precision of emotional response of the decoding paradigm, are relaxed. The aim of our *sonification paradigm* is to sonify an 'actual' emotion (or emotion complex), rather than to assume that a premeditated, single, simple, putative emotion is experienced.

We therefore conducted a study in which author N. F. chose extracts of film music that were intended to sonify each of six basic putative emotions (similar to the techniques used in the empirical studies cited above). Participants rated the emotions that the music expressed to see if the emotions selected were 'correctly' sonified. Our approach differs to much previous research as we do not approach the 'decoding' as an all or nothing selection, but as a complex response.

3. HYPOTHESIS AND DESIGN

Models of the structure of emotion can be broken into two broad groups: dimensional and discrete. Dimensional models assume that emotions can traverse from one to another (such as happy to sad), while discrete models assert that each emotion is a separate, independent entity. While the subtleties of the two models are more complex [23], we were interested, here, in allowing some degree of sophistication in rating response to those found in the literature and discussed above. For example, we wanted to be able to allow for multiple emotional responses to be reported simultaneously [22], and that they each be considered as part of the overall emotional response [24]. While retaining a reductionist, experimental approach, we wanted to increase the chance of identifying the actual emotion that was sonified. Our format was therefore a compromise between a dimensional model (of arousal and valence, the so-called two-dimensional emotion space, see [25]) and a discrete model. We selected 6 discrete emotions (that were relevant to music) that could be distributed in a more-or-less equidistant manner about the two dimensional emotion-space. The space consists of valence on the x-axis (positive emotions to the right, and negative to the left), and arousal on the y-axis (high arousal emotions for positive y and low arousal emotions for negative y values). This two dimensional model has been applied frequently in the past [25-27]. A mapping of basic emotions onto the two-dimensional emotion space are shown in Figure 1 using schematic facial expressions to represent the verbal labels used: Calm (bottom right quadrant), Happy (middle right), Excited (top right quadrant), Angry (top left quadrant), Scared (middle left) and Sad (bottom left quadrant). The schematic facial representation of these emotions shown in Figure 1 demonstrates the approximate geometric location of these emotions in emotion space, and provides a non-verbal way of rating the success of the sonification — although in the experiment described verbal labels were used.

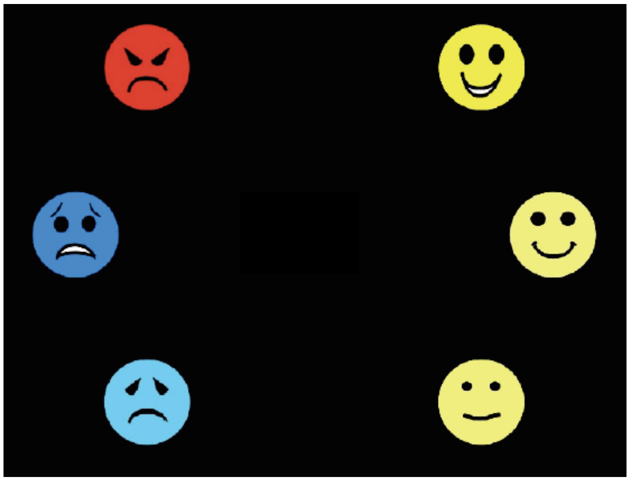

Figure 1: A schematic, face-based representation in two dimensional emotion space of the six selected emotions, as described in the main text.

While Scared does not clearly map to the middle left of the emotion space, it was selected because some previous research has demonstrated confusion in the communication of scared and angry emotions [13, see also 28 for other emotions that can get confused]. The benefit of our system, we propose, is that, like Hevner's, emotions that are close together around the two-dimensional circumplex [29] are more similar, and thus apparent 'miscodings' of the emotion sonified can be interpreted as similar if adjacent, rather than incorrect or confused.

We therefore hypothesise that participants will be able to detect the emotion that the music is expected to sonify, and we will measure the success of the sonification for each of the six discrete and distinct verbal emotion labels.

4. METHOD

4.1. Participants

Twenty-six students from a range of undergraduate and graduate music courses were invited to take part in the study in return for course credit. Participant ages ranged from 19 to 24 years, with 12 males and 14 females. Their self-reported years of instrumental lessons ranged from none to 18 years, with a mean of 11.0 years.

4.2. Stimuli

Stimuli were selected such that excerpts of film music were identified which might sonify each of the six 'target' emotions. Three examples of each were selected to reduce the chance of an eccentric example being selected by chance. The selection of music in this study was restricted to film music that did not have words. Film music has an advantage over several other musical styles for our purposes because it is explicitly programmatic and 'associationistic' [30]. Association, argues Cohen, "accounts for the direct transfer of meanings elicited by music to the film context, setting the mood, or disambiguating plot" (p. 29). The composers intention is, therefore, frequently to sonify the intended mood or emotion depicted by the scene (e.g. a thrilling car chase). Furthermore, these musical ideas need to be reasonably formulaic, so that the meanings associated with (and encoded into) the music can be clearly sonified. However, for a generally detectable sonification the music needed to have minimal personal, eccentric connections or associations—that is, not remind them of a personal, explicit situation. We restricted the selection of music to the soundtracks of a film genre that was likely to fulfill these criteria: Disney-Pixar fantasy-adventure animations (*Toy Story 3, A Bug's Life, Finding Nemo, Monsters Inc, Cars* and *Up*). With three excerpts of each target of emotion, 18 excerpts were selected ranging from 7 to 27 seconds in duration, as shown in *Table 1*. An additional 'Excited' excerpt was selected so that the total pool of excerpts was not a number that may have given away the number of excerpts in each target emotion category–the new total of 19 stimuli would reduce the likelihood that the participant divides 18 by 6 and uses a process of elimination strategies to select some responses, rather than responding to properties of the stimulus. The excerpts were edited so as to start and end as close to a phrase boundary as possible. The sound recordings were extracted for presentation to the participants in the same random order.

Stimulus Code	Film: Track Name	Start Time	Dur	Fam
Angry1	Up: 52 Chachki Pickup	00"53	17	5.42
Angry4	Toy Story 3: Come to Papa	00"38	20	4.85
Angry5	Toy Story 3: Cowboy!	03"36	19	4.65
Calm1	Finding Nemo: Wow	00"22	16	6.46
Calm2	Finding Nemo: Field Trip	00"00	21	5.23
Calm3	Finding Nemo: The Turtle Lope	00"48	20	5.96
Excited1	Toy Story: Infinity and Beyond	00"15	16	5.81
Excited4	Cars: The Piston Cup	00"05	7	5.38
Excited5	Cars: The Big Race	01"11	18	7.42
Excited3	Up: Memories Can Weigh You Down	00"26	21	6.54
Happy1	Cars: McQueen and Sally	00"04	16	5.81
Happy2	Monsters Inc.: Monsters, Inc.	00"06	15	6.77
Happy3	Up: Up with Titles	00"00	10	5.73
Sad1	Cars: Goodbye	00"00	27	6.27
Sad6	Toy Story 3: You Got Lucky	01"00	21	6.58
Sad7	Toy Story 3: So Long	02"20	23	5.85
Scared1	Cars: Mcqueen's Lost	00"55	11	6.04
Scared2	Up: The Explorer Motel	00"34	19	4.42
Scared4	Up: Giving Muntz the Bird	00"54	14	4.69

Table 1: Stimuli used in the Experiment. Three excerpts intended to sonify each emotion were selected, except for Excitement, which has four (see text for rationale). Stimuli are all instrumental, mostly orchestral arrangements. Familiarity (Fam) shows the mean rating for all participants on a scale of 0 (unfamiliar) to 10 (very familiar). Duration of excerpt (Dur) is in seconds. See Discography, §8, for details of sound recording sources.

4.3. Procedure

The study was designed and presented online using KeySurvey (http://www.keysurvey.com/) available via a URL link on the internet. After reading the ethics information, the participant was asked to be in a place that was as quiet and private as possible, and to wear headphones, if possible. They were then asked to click an icon to hear an excerpt, and then rate the excerpt in terms of the amount of each of the six emotions expressed on a scale of 0 (not at all) to 10 (a lot). They were also asked to rate how much they liked the excerpt and their familiarity (each also on a scale of 0 to 10, none to very respectively—see *Table 1*). When they completed responses to the 19 excerpts they were asked some demographic questions: languages spoken, age, gender and musical experience.

5. RESULTS AND DISCUSSION

The first analysis consisted of examining the number of times the highest scoring emotion matched the putative 'target' emotion. A single emotion could be selected up to 3 (excerpts) x 26 (participants) times (4 x 26 for Excited). *Table 2* demonstrates that the target emotion was successfully 'sonified' because the numbers along the diagonal are high (for example, the Calm excerpts were rated as highest on the Calm rating scale 56 times, compared to the other emotion ratings, but on 16 occasions, the same Calm stimuli were given Happy as the highest rating).

Highest Ratings Count ⇒ / Target Emotion ⇓	Angry	Calm	Excited	Happy	Sad	Scared
Angry	20	0	32	1	0	25
Calm	0	56	3	16	3	0
Excited	1	1	77	25	0	0
Happy	1	6	6	65	0	0
Sad	0	34	0	5	34	5
Scared	17	0	17	0	1	43

Table 2: Count of number of times the target emotion (first column) was rated highest of the six emotion scales.

From the decoding paradigm perspective we might conclude that some emotions were confused. Moving down the rows of *Table 2* we see Angry and Scared confused [row 1, and similar to the finding reported in 11]; Excited and Happy (row 3); Sad and Calm [row 5, and similar to 31]. Just as with previous studies [e.g., 10, 12, 13, 14] we find that music is not able to express the intended emotion 100% of the time.

However, the sonification paradigm treats emotion as a calibrated meter, where the sonified emotion (as distributed in emotion space, in Figure 1) better aligns with a point somewhere in between the putative 'target' emotion and an adjacent emotion. Just as a Geiger counter increases the click rate when more ionising radiation is detected, the sonified emotion 'meter' moves to somewhere in between two categories as an emotion that lies between Happy and Excited is indicated as such rather than one or the other. From this perspective, we see that in all cases when an emotion appears to be 'confused' by well represented emotion category ratings, it actually comes from a semantically nearby emotion rating. Our explanation, therefore, is that when the examples were selected, they were chosen to roughly evoke a putative emotion, and were not properly calibrated to the target.

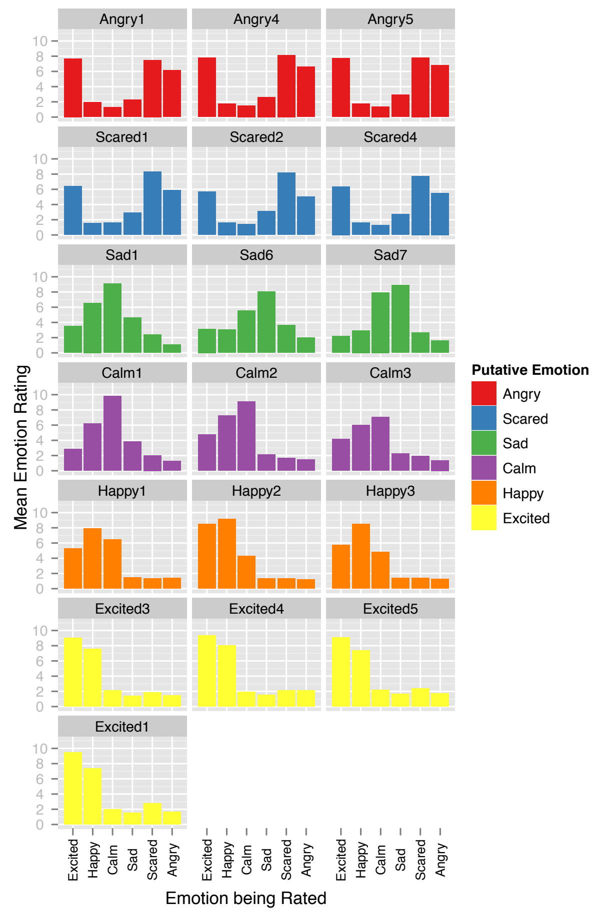

Figure 2- Results of experiment comparing the mean rating of 6 emotional categories for 19 musical excerpts. Excerpt codes (see Table 1) are presented at the top of each of the panes. Three exemplars of each target emotion (4 examplars for Excited) are plotted.

Consider the rating results for each stimulus, as shown in *Figure 2*. Here the emotions are shown in the sequence around the circumplex. In each case we see a peak in the rating at or near the target emotion. For example, Scared ratings for each of the three Scared target examples (top right pane) were high. However, Angry and Excited ratings were also high for all three of these cases. According to the decoding paradigm, Anger and Excited ratings would be considered confusions of Scared. But the sonification paradigm suggests that the target emotion was miscalibrated, and that the actual emotion sonified was some, possibly complex, combination of the three related emotions, or some averaging of them, perhaps pointing somewhere around Angry as the best estimate of the actual sonified emotion.

An obvious criticism of this interpretation is that anger and fear are qualitatively different emotions [32], making a mixture of them not necessarily meaningful. However, in all other cases, working down the triads of *Figure 2*, we see a meaningful relationship between emotional responses, since only adjacent categories to the target receive reasonably high ratings. Consider the four Excited targets. In each case Excited is rated highest, but also the Happy ratings are the next highest. Here we do not say that Happy and Excited emotions were confused (though the Lens model paradigm might suggest this). Instead, the sonified emotion is calibrated as being somewhere in between Excited and Happy, but slightly closer to Excited along the circumplex of the emotion space. We might still wish to make the same conclusion about Scared and Angry. Although they may be qualitatively different in important ways, there are still elements of valence (at least) that they do share, and are arguably as or more similar to each other than, for example, Angry and Sad.

Among the possible criticisms of our approach, we will now discuss two. One is on the question of the apparent 'upside-down' or circular nature of our approach to mapping emotions to sound. The other, partly in response to that issue, is a speculation about how emotion might be treated as data for sonification by musical fragments, more in line with some traditional, systematic approaches of sonification.

5.1. Emotion Sonification Mapping

Our approach collects data about emotional response to music by seeing how well it was sonified: Does piece A express (sonify) happiness better than it does calmness? Traditionally the variables to be sonified are the input data. Is our approach back to front? One key issue here is that converting emotion into data is non-trivial. Unlike variables such daily temperature readings, or distance of a surgeons knife from a target organ, emotion does not seem to reduce neatly into a variable that outputs one-dimensional numbers. This is, indeed, a reality for many factors in social and psychological sciences. Nevertheless, those fields continue to try to measure variables such as emotion, personality, cognition and behaviour. Suppose, then, that emotions could be reduced to numbers. One way this already happens is by using dimensional models of emotion, described earlier. We may have an emotion described as a series of numbers unfolding in time, with a time series for the valence component of the emotion and another simultaneous series for arousal. We then find variables to sonify those two sets of data. For example, valence might be coded very simply by mode (major for positive valence, and minor for negative valence), and simultaneously, arousal values are mapped onto tempo and loudness (high values of arousal map to fast tempo and loudness). This approach comes closer to satisfying Hermann's criteria [4]. In this case, however, we would need, as a listener, to understand that major mode represented a happy input, and so on. That is, we would need to learn the meaning represented by the sounds.

The argument in our paper is that we have learnt these meanings through our exposure to music and associations (mappings) through exposure within our culture(s). Social and psychological scientists are comfortable with the idea that the mapping will not be perfect, and that in fields of human measurement we are dealing with highly complex, multidimensional streams of data.

Since we are not yet privy to the full, data based coding of emotions (we do not yet know exactly what numbers, along which dimensions to place an emotion pertaining to sadness), we cannot define our sonification of emotion, yet, in a way compatible with Hermann's definition. Cultural factors, and individual difference add to the complexity, by providing additional noise to the underlying data signal. However, any signal can be susceptible to noise, and all are susceptible to measurement error.

A solution to this dilemma is to argue for a range of conceptualizations of sonification. Traditional approaches, such as those suggested by Hermann require highly specific stipulation of input parameter mapping to the auditory parameter(s). In sonifying emotions, this would mean streams of data representing time varying emotions, say along the arousal and valence dimensions, which are explicitly mapped to musical parameters, as described. In the conceptualisations of emotion sonification shown in Figure 3, this more traditional approach is indicated toward the bottom of the figure – 'useful for parameter mapping'. However, as we move up this spectrum of conceptualization, we relax these rules. In the present case, this makes the assumption that emotions are highly complex, and currently too complex to code easily, but also that music has already encoded these emotions through cultural exposure and biological factors. Supposedly the composer and/or performer is aware of these conventions and codings, but can convert emotions into music intuitively and holistically, without the need for explicit, reportable coding. The job of the sonification researcher then becomes to attempt to successfully reduce this process to its component parts, both in terms of emotional measurement, and musical parameter combinations that can code these emotions. We are not in a position, yet, to say for sure that this reduction would lead to a unique solution (e.g. that a specific emotion represented by happiness can be expressed uniquely and optimally by one set of musical features). And we do not know for sure whether individual differences and states within a given culture can simply be treated as measurement error. But, as discussed in this paper, music psychologists have been working on such matters.

Our approach, in the mean time, assumes that despite some variability humans already encode/decode emotions into/from music. Therefore measuring the output emotion conveyed by the music gives us some indication of the input emotion that has been coded. It is currently too complex to both quantify the emotion *and* find the mapping onto the multiple music parameters that sonify these emotions, and we are therefore forced to take an approach that is more complex for sonifying, as shown in the top part of Figure 3.

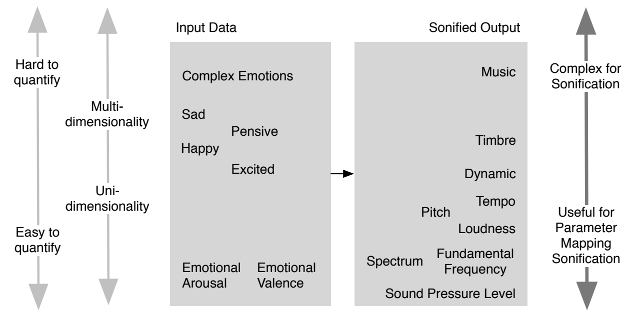

Figure 3 – Conceptualisations of emotion sonification.

5.2. Systematic Methods for Emotion Sonification Mapping

Given the complexity of traditional musical forms in sonifying emotions on the 'parameter mapping' end of the sonification conceptualisation spectrum, we speculate here on how computer generated or controlled music could be used. However, attempts to use computational generative music composition systems to drive the emotional content of music are likely to face similar barriers, despite the avoidance of the human element, as the emotional complex response is not known in advance. If we are actually interested in mapping arbitrary data directly to a listener's complex emotional response, rather than a purported musical or acoustical correlate, then collecting and *reflecting* emotional responses to musical stimuli used is required.

A possible method for achieving this is to extend the techniques developed for concatenative synthesis systems (such as Schwarz' Caterpillar [33, 34]) and use subjective emotional responses as an input for such an automated system to reorganize sound. Concatenative synthesis systems use time-aligned audio descriptors to describe an audio sample, in terms of for instance its sound pressure level (SPL) or pitch, and then divide an audio sample into small slices of audio that can then be rearranged. A simple example is to take a performance of a particular tune, extract each individual note, and then reorder the notes by their pitch in order to play a different tune. This process is usually only attempted with deterministic audio description algorithms (like pitch or SPL), but there is no theoretical impediment to using subjective continuous emotional responses in their place (perhaps obtained following Schubert [25]), as long as the time-alignment can be approximated with a reasonable degree of accuracy, and a dimensional model of musical emotion is used. Using this model of sonification of data to emotion, a set of listeners would have to emotionally respond to small slices of sound, and then these slices of sound would be rearranged based on a mapping of the ratings to the arbitrary input data to be sonified.

Practically, a method for sonification of data by emotional content of music is as follows:

1. Source music with a wide array of emotional content would be rated by a large number of respondents, resulting in numeric values of valence and arousal, either using reductionist approaches (emotional valence and arousal measures), or multi-dimensional emotion complex measures, for each audio time window being obtained.
2. These audio chunks are then annotated with the obtained data (perhaps the size of a beat or a bar of music).
3. Finally they are arranged in time based on their correspondence to the input data to be sonified, and concatenated to produce the sonification.

Ideally, the source music rated may be at a given tempo (or integer multiples), and use closely related musical keys, so that switching between segments with different emotional content results in greater musical coherence. This algorithm appears to satisfy Hermann's criteria for a sonification [4] – it represents the relationships in the input data; it is systematic and reproducible; and it can be undertaken with alternative input data.

An algorithm like the above would, however, be limited by the correspondence between the multi-dimensional emotional complex obtained from the musical stimulus and the format of the arbitrary data being sonified – which is in turn a function of the measuring instrument used to obtain the emotion complex. Other concerns that may arise with the use of emotional content of music as a data mapping target, are that music can have memory-related emotional effects, such as emotional contagion or episodic memory [21], rendering a listening to a sonification of emotion through music significantly influenced by, for example, the listener's past associations. Cultural effects are also difficult to control, as are time dependent ordering effects that may influence emotional responses, and the not inconsiderable effects of stimulus familiarity may further limit this type of sonification if it is repeated.

Although practically investigating this method is outside the scope of this paper, it does preface the possible applicability of the principles discussed and investigated outside the music-emotion domain and to arbitrary sonification contexts.

6. CONCLUSION

This paper discussed the idea that music may be treated as a culturally calibrated sonification of emotions. We reported an experiment where music was selected with the intention of sonifying (expressing) particular emotions. The results were interpreted in two ways. The traditional approach, which we classified as being based on a decoding paradigm (where an intended, putative emotion is encoded into the music and then decoded by the listener) resulted in some confusions of target emotions. This is because research published under such a paradigm is typically interested in which *discrete* emotion is *best* expressed by the music. However, by applying our paradigm of sonification, the results could be more adequately explained. Instead of expecting the premeditated, putative emotion to be identified, sonification treats the music as a 'measure' of emotion, and in so doing is more like an emotional meter that the listener hears and reports (the listener is still decoding, but the analysis is not dependent on the simplification of which single emotion is the best representation of the music. It is the trend across each rated emotion that is of interest). With this interpretation, the perceivers identify the emotions expressed, and we assume that the planned or intended putative emotional projections may have been miscalibrated. In the present study, the actual, composer/scenario 'intended' emotion expressed by the film music excerpt is better explained as a combination or mixture of several emotion categories distributed (in the present case) over an emotion space.

A corollary of this approach is that the measurement of the emotion expressed by music can be no more accurate than the measuring instrument. So, before the idea that emotions are not correctly communicated, or confused, are concluded, attention needs to be drawn to the measuring instrument. This is a situation peculiar to sonification of emotion, since most auditory displays convert numbers to sound. Deciphering the number displayed does not, typically, involve the same complexities in evaluation (that is, the need for an emotion measuring instrument), because the cultural conceptualization of number systems is considered objective and unambiguous, certainly when compared with emotions, which are more fuzzy – see Figure 3 and [35].

This means that the decoding paradigm is in fact redeemed. Our approach simply questions the way that the ideas emerging from decoding conceptualisations of musical emotions are typically found in practice. What we call a 'sonification' paradigm does fit with the way decoding paradigms are applied in the literature. As Juslin and others [28, 36] point out, these studies were generally interested in which discrete emotions music could *best* express, and that these were typically restricted to four (basic), categorical emotions, not assumed to be related to each other, contrary to our approach (six discrete emotions that could be aligned in emotion space). The sonification paradigm may, therefore, be seen as an extension and different interpretation of the decoding paradigm.

We also acknowledge that sometimes quite mixed and distinct emotions can be expressed by music, such as both happy and sad [22, 37, 38]. In response to this concern we make two responses. First, from a sonification perspective, it is a limitation of the measuring instrument that mixed emotions are not identified, rather than music being unable to sonify such different emotions. Further, it may be that other factors account for such complex interpretations—nostalgia, personal associations and so forth [21, 39]. This is why we argue that our study produced such 'clear' results. The measurement of the sonified emotions led to fairly simple emotional expressions (measurable by our instrument) because we deliberately tried to select music that was not overly familiar or unfamiliar (see familiarity scores by excerpt in *Table 1*) – reducing the likelihood of strong personal associations. For example, examining the ratings of the excerpt with the highest (Excited5, 7.42) and lowest (Scared2, 4.42) familiarity scores in *Figure 2*, the profiles are almost identical to the corresponding target emotion response profiles. Of course, an averaging effect, or regression to the mean, is one explanation for this similarity. But we assert that trying to select music that fits with cultural norms (in this case, through selection of instrumental, stereotyped film music, that is not overly familiar) may be one way of increasing the agreement as to the emotion that the music is sonifying. Future research should continue to test this assumption, and to examine music that may have more complex relations with individual listeners.

7. ACKNOWLEDGMENT

This research was supported by an Australian Research Council Discovery Project (DP1094998) awarded to the authors E. S. and G. M. We are grateful for the comments of an anonymous reviewer on an earlier version of this manuscript.

8. DISCOGRAPHY

Cars (Original Soundtrack); Released June 6, 2006; Walt Disney Records/Pixar 61349-7/Randy Newman

Disney Pixar Greatest Hits; Released 12 June, 2009; Walt Disney Records/Pixar 861290/Various

Finding Nemo (An Original Soundtrack); Released 20 May, 2003; Disney Enterprises, Inc./Pixar Animation Studios 60078-7/Thomas Newman & Bill Bernstein

Toy Story (Soundtrack from the Motion Picture); Released 1995; Buena Vista Pictures Distribution, Inc. 351041/ Chris Montan

Toy Story 3 (Music from the Motion Picture); Released 18 June, 2010; Walt Disney Records/Pixar 000470502/Chris Montan

Up (Original Soundtrack); Released 4 September, 2009; Walt Disney Records/Pixar (http://itunes.apple.com/us/album/up-soundtrack-from-motion/id316618105)/Michael Giacchino

9. REFERENCES

[1] G. Kramer, B. Walker, T. Bonebright, P. Cook, J. Flowers, N. Miner, J. Neuhoff and International Community for Auditory Display, *Sonification report: Status of the field and research agenda*: International Community for Auditory Display, 1999.

[2] S. Barrass and B. Zehner, "Responsive sonification of well-logs," in *Proceedings of the International Conference on Auditory Display*, Atlanta, GA, 2000, pp. 72–80.

[3] D. Worrall, M. Bylstra, S. Barrass, and R. Dean, "SoniPy: The design of an extendable software framework for sonification research and auditory display," in *Proc. of the 13th International Conference on Auditory Display*, Montréal, Canada, 2007.

[4] T. Hermann, "Taxonomy and definitions for sonification and auditory display," in *Proceedings of the 14th International Conference on Auditory Display*, Paris, 2008.

[5] P. Vickers and B. Hogg, "Sonification abstraite/sonification concrete: An 'aesthetic perspective space' for classifying auditory displays in the ars musica domain," in *Proceedings of the 12th International Conference on Auditory Display*, ed London, UK, 2006, pp. 210-216.

[6] K. Hevner, "Experimental studies of the elements of expression in music," *American Journal of Psychology,* vol. 48, pp. 246-268, 1936.

[7] E. P. Asmus, "The development of a multidimensional instrument for the measurement of affective responses to music.," *Psych. of Music,* vol. 13, pp. 19-30, 1985.

[8] C. L. Krumhansl, "An exploratory study of musical emotions and psychophysiology," *Canadian Journal of Exp. Psychology,* vol. 51, pp. 336-352, 1997.

[9] P. N. Juslin, "Five facets of musical expression: A psychologist's perspective on music performance," *Psychology of Music,* vol. 31, pp. 273-302, 2003.

[10] P. N. Juslin, "Cue utilization in communication of emotion in music performance: Relating performance to perception," *J. Exp. Psych.: Human Perception & Performance,* vol. 26, pp. 1797-1813, 2000.

[11] M. M. Terwogt and F. Van Grinsven, "Musical expression of moodstates," *Psychology of Music,* vol. 19, pp. 99-109, 1991.

[12] A. Gabrielsson and P. N. Juslin, "Emotional expression in music performance: Between the performer's intention and the listener's experience," *Psychology of Music,* vol. 24, pp. 68-91, 1996.

[13] P. N. Juslin, "Emotional communication in music performance: A functionalist perspective and some data," *Music Perception,* vol. 14, pp. 383-418, 1997.

[14] E. S. Nawrot, "The perception of emotional expression in music: Evidence from infants, children and adults," *Psychology of Music,* vol. 31, pp. 75-92, 2003.

[15] A. Gabrielsson and E. Lindström, "The role of structure in the musical expression of emotions," in *Handbook of Music and Emotion: Theory, Research, Applications*, P. N. Juslin and J. Sloboda, Eds., ed Oxford: OUP, 2010, pp. 367-399.

[16] S. K. Langer, *Philosophy in a new key: A study in the symbolism of reason, rite, and art.*, 3 ed. Cambridge, MA: Harvard University Press., 1957.

[17] L. Addis, *Of Mind and Music*. London: Cornell University Press, 1999.

[18] L.-L. Balkwill and W. F. Thompson, "A cross-cultural investigation of the perception of emotion in music: psychophysical and cultural cues," *Music Perception,* vol. 17, pp. 43-64, Fal 1999.

[19] S. Garrido and E. Schubert, "Imagination, empathy and dissociation in individual response to negative emotions in music," *Musica Humana,* vol. 2, pp. 53-78, 2010.

[20] S. Garrido and E. Schubert, "Individual Differences in the Enjoyment of Negative Emotion in Music: A Literature Review and Experiment," *Music Perception,* in press.

[21] P. N. Juslin and D. Västfjäll, "Emotional responses to music: The need to consider underlying mechanisms," *Behav. and Brain Sci.,* vol. 31, pp. 559-575, 2008.

[22] P. G. Hunter, E. G. Schellenberg, and U. Schimmack, "Feelings and perceptions of happiness and sadness induced by music: Similarities, differences, and mixed emotions," *Psychology of Aesthetics, Creativity, and the Arts,* vol. 4, p. 47, 2010.

[23] J. A. Sloboda and P. N. Juslin, "At the interface between the inner and outer world: Psychological perspectives," *Handbook of Music and Emotion: Theory, research, applications,* pp. 73-97, 2010.

[24] P. G. Hunter and E. G. Schellenberg, "Music and emotion," in *Music perception*, M. R. Jones*, et al.*, Eds., ed New York: Springer, 2010, pp. 129-164.

[25] E. Schubert, "Measuring emotion continuously: Validity and reliability of the two-dimensional emotion-space," *Australian Journal of Psychology,* vol. 51, pp. 154-165, 1999.

[26] J. A. Russell, A. Weiss, and G. A. Mendelsohn, "Affect Grid: a Single-Item Scale of Pleasure and Arousal," *Journal of Personality and Social Psychology,* vol. 57, pp. 493-502, 1989.

[27] A. Gabrielsson and E. Lindström, "The Role of Structure in the Musical Expression of Emotions," in *Handbook of Music and Emotion*, P. Juslin and J. A. Sloboda, Eds., ed Oxford: OUP, 2010, pp. 367-400.

[28] P. N. Juslin and P. Laukka, "Communication of emotions in vocal expression and music performance: Different channels, same code?," *Psychological Bulletin,* vol. 129, pp. 770-814, Sep 2003.

[29] J. A. Russell, "A circumplex model of affect," *Journal of Social Psychology,* vol. 39, pp. 1161-1178, 1980.

[30] A. Cohen, "How music influences the interpretation of film and video: Approaches from experimental psychology," *Selected Reports in Ethnomusicology: Perspectives in Systematic Musicology,* vol. 12, pp. 15-36, 2005.

[31] N. Gosselin, I. Peretz, E. Johnsen, and R. Adolphs, "Amygdala damage impairs emotion recognition from music," *Neuropsychologia,* vol. 45, pp. 236-244, 2007.

[32] A. Marsh, N. Ambady, and R. Kleck, "The effects of fear and anger facial expressions on approach-and avoidance-related behaviors," *Emotion,* vol. 5, pp. 119-124, 2005.

[33] D. Schwarz, "The caterpillar system for data-driven concatenative sound synthesis," in *6th International Conference on Digital Audio Effects*, London, UK, 2003.

[34] D. Schwarz, R. Cahen, and S. Britton, "Principles and Applications of Interactive Corpus-based Concatenative Synthesis," in *Journées d'Informatique Musicale (JIM'08)*, Albi, 2008.

[35] L. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning—I," *Information sciences,* vol. 8, pp. 199-249, 1975.

[36] J. E. Resnicow, P. Salovey, and B. H. Repp, "Is recognition of emotion in music performance an aspect of emotional intelligence?," *Music Perception,* vol. 22, pp. 145-158, 2004.

[37] P. G. Hunter, E. G. Schellenberg, and U. Schimmack, "Mixed affective responses to music with conflicting cues," *Cognition & Emotion,* vol. 22, pp. 327-352, 2008.

[38] P. Evans and E. Schubert, "Relationships Between Expressed and Felt Emotions in Music," *Musicae Scientiae,* vol. 12, pp. 75-99, 2008.

[39] F. Barrett, K. Grimm, R. Robins, T. Wildschut, C. Sedikides, and P. Janata, "Music-evoked nostalgia: affect, memory, and personality," *Emotion,* vol. 10, p. 390, 2010.