A SIGNAL-FLOW MODEL FOR THE DESIGN OF SOCIAL BIOFEEDBACK SYSTEMS

by

RONALD GRAEME McCAIG

B.A.Sc., The University of British Columbia, 2001

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

(Electrical and Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA

August 2006

Abstract

This paper explores the design of 'social biofeedback systems' (SBFS's) by presenting a theoretical 'signal-flow model' and a particular system design called *2Hearts.* A proposed definition is given of SBFS's as systems which sense and display biosensor-derived data in real-time, involve multiple people in social-communication, and convey information relative to emotion or affect. Previous SBFS systems and other related works are reviewed to show both the pattern of interest in such systems and the lack of an underlying theoretical framework.

The signal-flow model is developed in detail with reference to literature from psychology, auditory display and affective computing; in its current form the model is focused on SBFS's which employ heart-sensor data as input and audio as output. Key features of the model are a distinction between conscious and unconscious perception/ processing of emotional information, and identification of several different representations of affect information within the brain that are useful for SBFS's to replicate or process. Suggested design tools are described including SBFS design-goal 'paradigms' and a high-level SBFS design procedure, framed in terms of the signal-flow model.

The design and implementation of the 2Hearts system is described and justified according to the signal-flow model. The 2Hearts system incorporates time-compression/looping techniques to display different heart-data-derived variables in a synchronized, layered fashion, leading to the perception of rhythmic, musical audio patterns. The design and results of 2Hearts user testing are described. Significant testing results are obtained which characterize the emotional associations of some of the 2Hearts audio-display mappings, and lead to suggestions for improvement of these mappings. Other significant results suggest that people can voluntarily influence each others' heart rates using communicative strategies, thus validating one important assumption of SBFS design. Recommendations are given for future development of the 2Hearts system, the signal-flow model, and the user testing strategies.

Table of Contents

List of Tables

List of Figures

1. Introduction

This thesis presents a theoretical framework and a signal-flow model for the design of Social Biofeedback Systems (SBFS's). It also presents the design and testing of the example system, "2Hearts", which was used to develop and validate the theory.

1.1 Background and Motivation

Social biofeedback systems are currently a small research topic in the field of human-computer interaction. The term itself, *social biofeedback system*, is relatively novel. The term is proposed here as a suitable description for interactive, electronic systems with the following characteristics:

Characteristics Defining Social Biofeedback Systems

1. SBFS's take inputs and provide outputs to two or more human users, who are engaged in social interaction with each other.
2. Sensor inputs to SBFS's include physiological signals connected to psychological variables, (i.e. psychophysiological signals), such as heartbeat, breath, temperature, skin conductance, EEG (brainwave), etc.
3. Outputs from SBFS's display to the users information derived from the physiological signals and believed to be relevant to the emotional/social context.
4. SBFS's function at least partially in real-time (in the soft or hard sense of real-time).
5. SBFS tend to be associated with the detection/display of emotion or affect.

In connection with point 5, most well-studied physiological signals associated with psychological phenomena are related to the emotional or *affective* side of human experience. In social interaction, cognitive/rational factors often make up the primary topic of explicit, verbal communication. Thus a theme of SBFS's is often the revelation of the implicit or hidden.

Motivations for creating SBFS's include enhancing the effectiveness/intimacy of human social interaction by allowing users to better understand each others' psychophysiological reactions, and creating unique socially-controlled electronic experiences for entertainment or education. Some examples of potential applications are:

Potential Applications of SBFS Technology

- Remote communications-computer(chat/conference), augmented cellphones
- Home entertainment- music/visual art, video games, game music
- Psychological Display System for therapy (w/therapist, family therapy), spiritual exercise aid (yoga teaching etc.)
- Entertainment at Clubs etc. group audio/visual manipulation via conversation, touch, dancing
- Systems for Physical Exercise motivation, sports training etc.
- Performance enhancement for social working relationships at the office, meetings, etc. (a future possibility if systems become widely accepted)

SBFS's are also important as representative of a recent trend or possibility in the development of computer systems; the creation of systems which engage with 'soft' affective/emotional information. Most current uses of computers to enhance human communication involve the manipulation of 'hard', verbal/rational data- examples include word processing and math calculation. The processing of human emotional data is more complex and less well-understood, and thus such 'soft' technology is still in its early stages. The topic of "Affective Computing", discussed further in Section 2.3, has been established by Picard and other researchers to address this type of technology [1].

A final benefit to studying SBFS's is that a detailed analysis of SBFS's design and operation requires the integration of a wide body of literature from a multitude of disciplines, including Psychology, Fine Arts, and Engineering. This effort identifies many interesting themes and connections between the various research areas.

1.2 Objectives, Scope and Approach

This thesis project has evolved in its objectives since its beginning. The initial goal was to create a particular type of SBFS: a two-person musical instrument, with heartbeat sensors as the input. This system concept, given the working name "2Hearts", has remained a focus of investigation. However, once initial pilot systems were explored and the conceptual challenges of SBFS's were discovered, it became evident that developing a stronger theoretical basis for SBFS's would facilitate the progress of 2Hearts and also provide a useful research contribution to this emerging field.

The theoretical work has consisted of the development of a *signal-flow model* for SBFS design. The model continues to be developed as a useful theoretical concept which can apply to SBFS's in general. However, because 2Hearts was used as the initial example system, *the model is currently slanted toward describing SBFS's which use heartbeat sensors as input, and audio/music as the output.* There is a special attraction to exploring this input/output pairing because (as will be shown) heartbeats and music both incorporate *rhythmic patterns which indicate or convey affective information.* This overarching theme will be discussed at many points in the thesis.

Three general goals of the thesis project were:

- Development of a Signal-Flow Model for SBFS's
- Development of exploratory system 2Hearts, which aided in developing the model
- User testing of 2Hearts to validate the system and the model, and to supply additional data for developing the model

In more detail, the objectives for the signal-flow model were:

- Allowing the diverse body of research relevant to SBFS's to be organized and synthesized
- Providing a framework for discussion of SBFS designs by labeling important components and pathways of influence
- Simplifying reasoning about SBFS design/analysis by identifying pathways that can be conceptually compressed or ignored
- Highlighting missing areas of knowledge to address with future research
- Facilitating the definition of paradigms within which the objectives of a particular SBFS can be categorized
- Suggesting recommendations/predictions about how to properly design a SBFS to achieve given aims
- Suggesting appropriate testing strategies to measure the performance of SBFS's

2. Related Work

This section gives an overview of previous work in two areas related to the goals of the present project:

1. Actual SBFS's, or systems sharing some important attributes of a SBFS
2. Current theoretical frameworks which address some of the areas covered by the proposed SBFS signal-flow model.

2.1 Related Systems

This section describes a number of previous research system and artworks which meet the SBFS definition above or partially meet it. These works are relevant to the current project in several ways:

- They suggest theoretical perspectives and implementation strategies useful for SBFS design
- They provide evidence of a broad-based opinion or instinct that SBFS technology can be useful or aesthetically enriching.
- They illustrate the relative lack of a detailed theoretical model for general SBFS design/analysis.

Portable Gold and Philosophers' Stones **(1972)**

This electronic music work utilized 4 performers equipped with EEG, skin conductance and temperature sensors. Sensor data was filtered and processed in a number of ways with analogue and digital circuitry to produce a modulated drone-scape of audio [2]. It is suggested that this work meets the SBFS definition since performance of a musical quartet constitutes a form of social interaction (although performers may not be verbally interacting).

This work was one of several pioneering works using biodata to control music, demonstrating that the attraction of linking biosignal data to music has been observed for quite some time. Also, its emphasis on dividing ranges of biosignal behaviour using filter banks has been continued in the 2Hearts system design.

Biomuse (1992)

Not a SBFS, but rather a related technological tool, this commercial product is a sensor/processing suite designed for producing biosignal-controlled music and art [3]. Specific works of interest using the *Biomuse* have not been identified, but it is likely that works with this system have produced some valuable findings regarding compelling biosignal-music mappings.

Audible Distance (1997)

In this interactive art installation, multiple participants each wear ear-clip heartbeat sensors, and move around in a common area while wearing head-mounted displays which replace their normal vision with a simple virtual space [4]. Each participant is depicted in the virtual space as a simple pulsing globe which makes a heartbeat sound as it pulses. The pulses and noises are synchronized with each participant's detected heartbeats, and grow larger/louder as one participant moves physically closer to the other. *Audible Distance* clearly meets the definition of SBFS given the use of biosensor, audio-visual display and focus on user-user interaction.

The artist's description of the theoretical intent of the piece [4] dwells on the theme of connecting psychological distance with physical distance. The importance of heartbeat display is mainly left implicit; but perhaps heartbeats were though to be a strong symbol of the concept of 'self' and a simple but direct indicator of social arousal or affect.

Consideration of *Audible Distance* was useful to the current research by highlighting the fact the SBFS may work more effectively for some purposes by actually damping normal social communication channels.

Heartbeat: mb (1998)

In this dance performance, the famous dancer Baryshnikov was equipped with a wireless heart sensor, and danced to a musical background based on heartbeat sounds triggered by the sensor data [5]. The piece was said to be popular with audiences due to creating an intense connection with the performer's extreme physical exertion. This could be classified as an SBFS if the performer/audience relationship is considered to be a form of social interaction.

Unconscious Flow (1999)

This interactive art installation is a clear example of a SBFS: two users interact socially while equipped with heart sensors. Heart data is processed into 2 parameters, 'relaxation/strain' and 'interest', which are used to control the behaviour of 2 graphical computer-generated 'mermaid' characters [6].

Thus, the mermaids are supposed to mirror the 'true' implicit state of the emotional inter-user connection. The work's discussed assertion is that normally-hidden aspects of social relationship can be revealed by biosignals and employed beneficially for increased interpersonal understanding.

Heart Rate Sonification: A New Approach to Cardiopulmonary Diagnosis (2000)

This research project explores offline, time-compressed audio display of heart rate data [7]. The purpose of display is to enhance the perception of medically relevant heart rate variability (HRV) patterns. This work is not considered an SBFS due to the lack of real-time and social elements.

The background to this work is the research field of auditory display. The insights obtainable from the work are that audio display can uncover previously obscure patterns in biosignal data; and that compressing the data display to different time-scales can change the nature and accuracy of its perception. The *HRV Sonification* project diverges from the present work in that the affective aspect of heart rate and of audio display were not primary considerations.

Galvactivator / Conductive Chat (2001 / 2002)

The *Galvactivator* is a light-weight 'wearable' affective communication/augmentation device- a glove which combines a skin conductance sensor with a red LED which grows bright during episodes of sensed arousal [8]. Intended to let users explore arousal effects in social communication situation, the *Galvactivator* is certainly a SBFS.

Conductive Chat is a SBFS built on *Galvactivator* technology, which explores the remote-communication context by changing the visual appearance (color and size) of text in a computer 'chat' application according to each user's arousal level at the moment of typing [9].

Several findings from these projects are of interest to the current SBFS research:

- Interesting episodes of synchronized biosignal behaviour between remote users suggests the use of synchronization measures as indices of social relationship.
- Biosignal display used by crowds of people *en masse* can indicate relevant emotional trends such as boredom or anticipation.
- These projects have involved visual displays designed to present correct instinctive parameter associations (e.g. red = aroused). Musical display used in the current project was found to crucially depend on such correct associations, with major consequences resulting from good/poor design choices.

Pimp My Heart (2006)

This art installation combines a customized cruiser automobile and high-power audio amplification system with an earclip-based heart sensor and laptop-pc processing unit [10]. The heartbeats of the driver's seat occupant are either amplified directly (to onlookers and to the driver through haptic feedback) or used to modulate the playback speed of digital music files. It is suggested that *Pimp My Heart* meets the definition of SBFS, since driver and onlookers are connected in a particular asymmetrical social communication.

Theoretical discussion of the work refers to the display of heart data "as achieving an ultimate unity between car and driver" [10] and "addressing the vulnerability of human body and emotion" [10]. This work emphasizes the supposed link between heart, emotion, and identity, without any particular reference to theory.

The Affective Remixer (2006)

This research system uses skin conductance and foot-tapping sensors to infer the emotional state of the user, then selects audio segments from a matrix of synthetically – generated music clips, so as to influence the user's mood toward desirable state [11]. It is not a SBFS because it focuses on one isolated user.

The theoretical approach behind *Affective Remixer* involves parameterizing both music and sensor data into respective 2-dimensional spaces and determining affective state by using statistical methods (a Markov chain) on biodata. This approach exemplifies a developing trend, affective systems which employ statistical, categorical models of user emotional- and goal-state, and take discrete actions in reaction to such states. This perspective is currently the closest candidate for a commonly-used theoretical model for SBFS design. However, it does not fully describe many of the other systems described in this section.

It will be shown in this paper that the described approach is important, but not the only method for transmitting/manipulating affect-related information. The SBFS signal-flow model has the potential to place such approaches within a larger space of possibilities.

Bio-Informatic Feedback (2006)

This research system is in some ways similar to the *Affective Remixer*, but involves an instrumentalist user who actively plays music from a dynamically generated music score [12]. Skin conductance sensors on the performer lead to scaled 'activity' values which provide input to a real-time selection process of pre-composed score 'cells'. If performed for an audience, this system could be called a SBFS since the performer/audience relationship is partly social.

The theory behind the work makes mention of 'excited' and 'less excited' emotional states, which can be influenced by playing/hearing music- a basic assumption which is explored in detail by the SBFS model.

As seen from these examples, the idea of augmenting human communication in a deliberate way to change the nature of the social interaction has appealed to more than one researcher; however most systems developed so far have been essentially proof-of-concept.

Design has tended to begin with a specific approach to making the mapping between biosensor signals and audiovisual display, and accomplishing the artistic and engineering design necessary to realize this approach. Evaluation of the systems ranged from purely artistic (as in *Unconscious Flow* or *Audible Distance*) to statistics quantifying usage patterns and signal synchronization (as in *Conductive Chat*).

It is suggested that some developments needed to advance this field are:

- an agreed upon set of goals or specific applications
- standard 'best' tools for given applications (such as biosensor types and signal processing algorithms)
- accepted testing paradigms or standard measures for system effectiveness
- models for the action of the overall system (including the humans' interaction and the technology itself)

This project addresses these points through the development of the SBFS signal-flow model.

2.2 Related Theoretical Frameworks

Biofeedback

Social biofeedback systems have developed partly from the simpler concept of biofeedback. Although related to ancient techniques such as meditation, the term and concept 'biofeedback' was popularized in the late 60's and 70's. The aim of biofeedback is to provide real-time information about a user's physiological signals back to the conscious mind of that user. Maintaining and practising this link can allow the user to bring the physiological phenomena under voluntary control [13]. This principle is a valuable finding to be applied to the study of SBFS's. The biofeedback literature could be further researched for relevant findings on specific feedback modalities and psychophysiological phenomena. Social biofeedback systems expand on the biofeedback idea by increasing the number of users to two or more- the feedback paths then include intra-user feedback and inter-user feedback.

Affective Computing

A well-defined theory of SBFS design must include description of human interaction, human emotion, biosignal action and system components. Research incorporating each of these factors simultaneously has not been found to be common. The best example of an existing theoretical framework applicable to SBFS design can be found in the recently growing area of Affective Computing.

Within the field of Human-Computer Interaction is an area of research known as Affective Computing, defined as "computing that relates to, arises from, or deliberately influences emotions" [1:3]. Sub-areas of Affective Computing include 'affective wearables' and affect augmentation. Picard's book *Affective Computing* lays the groundwork for much of the current development in these areas. An example of Picard discussing the SBFS goal (in the context of virtual environments) is:

> "[...] communication through virtual environments could provide new channels for affect—perhaps, as one idea, via sensors that detect physiological information and relay its significant information. In this way, computer-mediated communication might potentially have *higher* affective bandwidth than traditional 'in person' communication." [1:57]

Relevant models presented/discussed by Picard in the book include:

- Modeling an Affective [Computer] System
- (HMM) Model for Affective Behaviour
- Various models of the cognitive structure of emotions

Much of this information is relevant to SBFS design. However, none of these models function at the particular level of detail developed here in the signal-flow model, and none of them integrate all of the various pathways present in a SBFS. Two factors in particular distinguish the signal-flow model from *Affective Computing* and the related literature:

1. The signal-flow model describes, within one model, inner and outer connections between two or more socially interacting users and an artificial system linked to them; all at a similar medium-level functional perspective useful to system design. Other Affective Computing models have explicitly map only part of this complete cycle at a time.
2. The signal-flow model distinguishes between categorical, dimensional and contour-based theories of emotion while locating aspects of each viewpoint within the SBFS signal-flow space. Most other Affective Computing work tends either to focus on a categorical/probability-based emotion model, or use direct/contour-based communication ideas in a more informal way.

Music and Communication

It is much easier to find in the literature theoretical foundations for individual pieces of the SBFS context. Partial models covering different portions of SBFS's have been noticed in the literature of psychophysiology and auditory display. A particularly relevant example is the Extended Lens Model of Juslin and Lindstrom which analyzes the communication of emotion in music between composer, performer and listener [14]. Being focused on traditional musical communication, this work naturally lacks certain elements of the signal-flow model:

1. The Extended Lens Model does not frame musical communication in terms of display parameters for a digital system
2. The Extended Lens Model does not deal with the emotion-biosignal link.

The theoretical models of Affective Computing and music-communication research remain very useful and are drawn upon further in the development of the signal-flow model presented here.

3. The Signal-Flow Model: Introduction and Initial Development

This chapter introduces the *SBFS signal-flow model,* a high-level model of the important interactions that occur in the functioning of a social biofeedback system, including inter-human interactions, intra-human interactions, and human-computer interaction. The signal-flow model was developed over time by combining literature findings with experience from the design of trial/prototype systems.

In line with the objectives listed in Section 1.2, the model is intended to simplify the complex network of elements and connections occurring in human social interaction and SBFS interaction, and to emphasize key design considerations relating to information flow in the SBFS.

3.1 Elements and Connections in the Signal-Flow Model

As introduced, the signal-flow model consists of nodes, which identify important conceptual units of information processing/storage, and arcs, which indicate information flow between the nodes. This section briefly describes each of the basic nodes and arcs employed in the model, in order to provide an overview. The identification and separation of nodes is then justified according to results from the literature cited in following chapters, and further detail is added to the basic model.

Figure 3-1 shows the SBFS signal-flow model. Note that the model is presented for the case of two users, but it can be readily extended to 3 or more users.

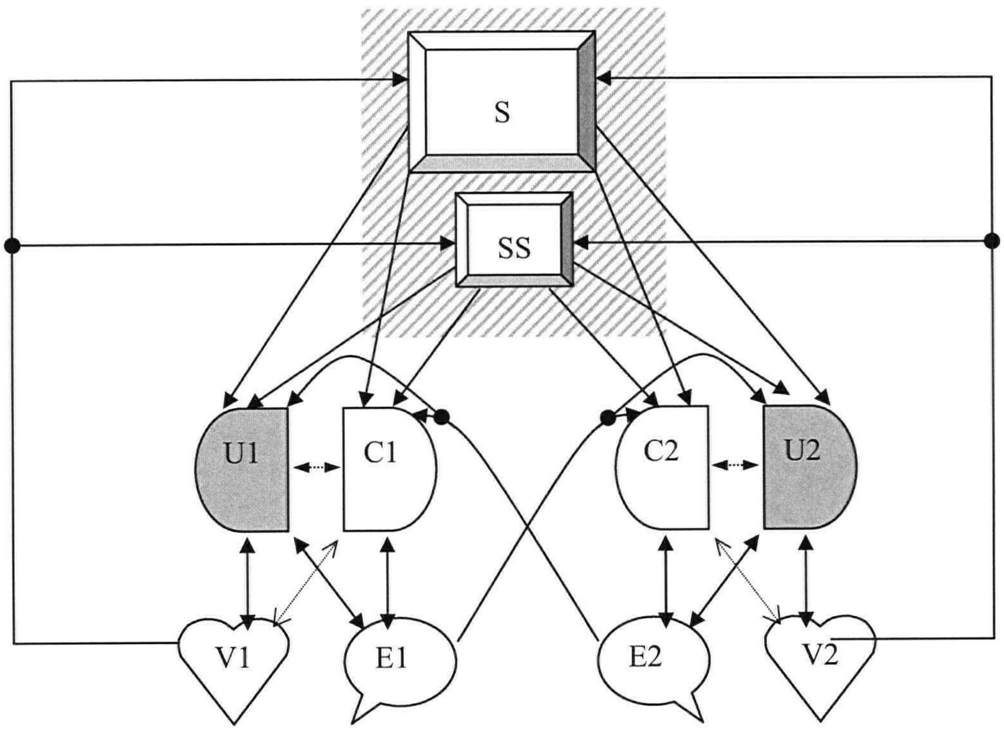


Figure 3-1. Basic Signal-Flow Model of Social Biofeedback. Includes the participants and the system. Thickness of arc indicates strength/importance of connection. Nodes are defined and discussed in the text--briefly: U= Unconscious, C= Conscious, V= Visceral, E = External (Communication), S= System (Per-User Processing) and SS= System (Joint User Processing)

Explanation of the Nodes

U1, U2: The **U**nconscious portion of user 1/user 2's brain. In this model the unconscious is defined as any brain processing activity of which the user is not aware. Thus this unconscious is the processor of low-level perceptual information, and the agent of control of automatic processes, such as the secretion of hormones and control of the digestive tract. Some processes, such as breathing, can shift between conscious and unconscious control depending on the user's focus of attention. In this model the unconscious is closely tied to the storage and processing of affective/emotional information.

C1, C2: The **C**onscious part of user 1/user 2's brain; the aspect of brain functioning that is qualitatively experienced including rational/verbal thought.

The conscious/unconscious distinction is emphasized in the signal flow model for several reasons. For one reason, the unconscious mind is the location of many of the processes that constitute what we regard as emotion- there is evidence that our conscious experience of emotion, which could be called "feeling", is only a small part of the picture [15].This is discussed in Chapter 4. Another reason to distinguish the conscious from the unconscious is to properly analyze the effects of audio and visual feedback. Different parameters and symbols in sound or image are capable of being perceived both consciously and unconsciously, and of having differing effects on the conscious and the unconscious mind. This is discussed further in Chapters 4 and 6. In the signal-flow model, the heart is controlled primarily by the unconscious, although this can be changed to some degree by the biofeedback mechanism.

V1, V2: The **V**isceral/somatic actions of user 1/user 2's body- defined as actions of the organs that are difficult for an outside observer to perceive, such as digestion, heartbeat, and electro-dermal skin response (EDR), and hidden movement/tension in the non-visceral (somatic) muscles.

E1, E2: The **E**xternal actions of user 1/user 2's body- actions that are perceptible to an outside observer, such as talking, gesturing, etc. External actions include explicit communication, such as verbal content, and implicit or non-verbal communication, such as tone of voice. Somatic actions which **are** perceivable to an external observer are included here.

The distinction between visceral and external in this model is one of convenience- in actuality this is another continuum, with actions such as breathing occupying a midpoint between, for example, speech and the action of the pancreas. But the two groupings provide a convenient perspective.

S: The artificial System which enhances social biofeedback by detecting users' visceral signals via biosensors, and transmitting information to the users via audio and/or visual (or perhaps other modalities such as tactile or olfactory) display.

SS: The user-user joint correlation/synchronization node. This node is shown in the model as separate from S, to emphasize the importance of inter-user visceral phenomena, processed separately from per-user phenomena. Such phenomena are discussed in Chapter 5. (Note that it is assumed that in natural, unaugmented social interaction, person-person visceral correlation is not usually noticed or processed by the people involved.)

The S and SS nodes are defined here in a way intended to be most helpful to the SBFS designer. It would be possible to show two S nodes, each taking input from one user; however, the simpler diagram suggested here has been found to be adequate for describing system function.

Explanation of the Arcs and Units of Measurement

Just as a traditional control-system model can incorporate translations between different forms of signal and energy, the Social Biofeedback signal-flow model contains arcs which represent a variety of signal types.

When discussing arcs and pathways in the signal-flow model, the discussion of a pathway from User 1 to User 2 (for example, E1→C2) will apply equally to the corresponding pathway (e.g. E2→C1), unless stated otherwise. Also, nodes without indices will be used to refer to the nodes of any one user. (For example, C→E implies a conscious to external connection *within* one user.) A notation such as S→Un is used to refer to both S→U1 and S→U2 at once.

Arcs interconnecting node types U, C, V, and E within one user's body represent a variety of neurological and physiological processes. Though it may be possible in theory to define exact units for some processes in terms of electrical nerve firing action, it is sufficient for many purposes to think of these signals qualitatively without specified units. A crucial assumption of this model is that *the path from U→V is considered stronger than the path from C→V, and also usually stronger than the combined path C→U→V.* This results from the fact that paths to the unconscious are defined here as stimuli that can be processed without conscious awareness. Thus, to give an example, if we give User 1 a message in Morse Code and he does not pay attention to translating the message, it is unlikely to affect his heart rate. If User 1 consciously translates the message, it may affect his heart rate through path C→U→V, depending on its content. However, if User 1 is afraid of spiders, viewing a spider will stimulate his unconscious as well as conscious mind directly, and therefore have a greater affect on the heart rate. This aspect of the model will be supported by current brain research in Chapter 4.

Arcs Vn→S represent human bio-signals as measured by sensors. Since these signals are in most cases digitized by the system, these arcs could be considered in terms of sampled digital information. For example, EDR could be considered as a time series of conductance values in microSiemens. However, these arcs will more often be considered in terms of higher-level variables derived from the bio-signals: for example a heart signal considered in terms of inter-beat interval (see Chapter 4).

Arcs directly connecting the users (e.g. E1→C2) represent physical actions that are transmitted through light, sound, touch etc. and then perceived through the sense organs- these signals must be considered in a variety of ways including symbolic content and modulation parameters (for instance speech should be analyzed for semantic content as well as tonal qualities.) For systems such as 2Hearts which attempt to augment the natural social context without disturbing normal communication channels, we need only consider the changes it makes to these arcs. Comparing systems with completely different interaction environments (e.g. comparing a face-to-face system with a remote chat system such as *Conductive Chat* [9]) is challenging because the system-influenced actions of these arcs must be analyzed and included in the comparison.

Arcs extending from S and SS represent display signals created by the computerized social biofeedback system. These can take the form of digital audio or video streams, however it is most useful to consider these signals as a time-series of higher-level parameter changes according to the scheme for sound/graphics generation within the computer system. For example, sound displays should be considered in terms of whichever parameters are modulated by the system (note onset, pitch, volume, timbre selection, etc.; see Chapter 6.) Later in the text, it will be shown that these parameters should be compared and characterized taking into account perceptual effects (see Chapter 6). *As the model is developed below, certain perceptual phenomena will be taken out of the "black box" of the mind and considered as part of the system to human-internal-representation pathway.*

Arcs U1→E1 and U2→E2 represent both external actions that are performed involuntarily (e.g. some facial expressions and gestures) intentional actions that are modulated unintentionally (e.g. agitated quavering in the voice); see Chapter 5 for further discussion.

3.2 Feedback in the Signal-Flow Model: An Initial Examination

One benefit of using a signal-flow type of analysis is that feedback phenomena can be identified and analyzed. Several types of feedback phenomena are important in the SBFS signal-flow model: naturally occurring feedback, control-enhancing artificial feedback, and the potential of unstable feedback loops.

In the model, several arcs are shown as double-headed arrows. This reflects the simplifying viewpoint that the brain mechanisms devoted to controlling the viscera act as two-way channels, a form of natural biofeedback. For instance, the nervous system connections between heart and brain include both "efferent" ("outgoing", brain to heart) and "afferent" ("incoming", heart to brain) nerves [16].

From a general control systems point of view, a control unit can often operate more effectively if provided with information about the effects of its control signals on the system (closed-loop control). The traditional idea of biofeedback could be viewed as the augmentation of closed-loop control between the conscious (controller) and the viscera (object of control). By creating the path V→C, control in the direction C→V is enhanced. Notably this enhancement can be partially retained once the system is removed, due to learning effects. A key hypothesis for SBFS's is that this enhancement effect can be generalized to a multi-person scenario: by creating the pathway U1→V1→S→C2, User 2 may gain an enhanced ability to consciously control User 1's emotions. Similarly, the pathway U1→V1→S→U2 may facilitate User 2 to unconsciously control User 1's emotions.

Feedback loops must also be considered for potential unstable behaviour. It can be seen that SBFS's give the possibility of positive feedback loops, for example U1→V1→S→U1. This loop could be predicted to exaggerate changes in the visceral behaviour. Picard points out that such exaggerations could in theory be physiologically hazardous, especially if the feedback is too intense, but characterizes music/audio as a display modality which is likely to be safe [1]. In the present research, this safety concern has been addressed by closely monitoring users at all times to watch for unstable visceral signals. Significant examples of such phenomena have not been observed.

3.3 System Goal/Usage Paradigms for Social Biofeedback Systems

Although the nodes and arcs of the signal-flow model will be analyzed in more detail in following chapters, even the basic model (in fact, a simplification of the basic model), along with considerations of feedback, can be used helpfully for the purposes of SBFS taxonomy and clarification of system design goals.

This section proposes the concept of *social biofeedback paradigms,* categories of SBFS function that can be described as patterns of information flow in the signal-flow model. The paradigms are distinguished based on *the cumulative information flow between the unconscious and conscious mind nodes.* Diagrams can be drawn based on this principle, including only the nodes Un and Cn, and drawing arcs between them to represent the cumulative pathways between these four nodes.

Also included is an elaboration of the arc representation: Arcs can be labeled "+" or "-" depending on whether the net action of the artificial system **S** is to enhance, or weaken that pathway. Neutral arcs (those unchanged by the action of the artificial system) can be left out altogether. (Note that a "+" pathway is a **strong** pathway, capable of effectively conveying either positive or negative affect information; a "-" pathway is **weaker** and will not convey positive or negative affect information as effectively.)

(One-Person) Biofeedback

As discussed in Section 3.2, this paradigm conveys the visceral signals of one user to the conscious mind of that user, with the aim of increasing understanding and control of the user's visceral and unconscious actions. Figure 3-2 shows this paradigm.

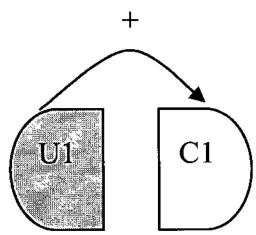


Figure 3-2. Biofeedback SBFS Paradigm.

Perform/Influence

This paradigm has the goal of influencing the emotional state of one user (the audience) to correspond with the emotional state of another user (the performer). A positive feedback loop can be applied to the performer's unconscious signals in order to create more dynamic unconscious changes on the part of the performer. This paradigm exemplifies one of the ideals of artistic performance: "self-expression" leading to extreme fellow-feeling.

An example of this paradigm is *Heartbeat:mb* [5] in which audience members were influenced to feel powerfully affected by the rise and fall of Baryshnikov's exertion level. Figure 3-3 shows this paradigm.

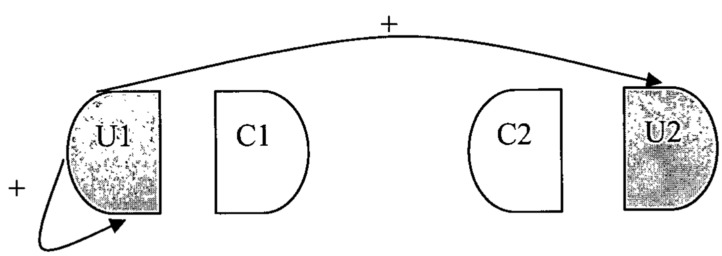


Figure 3-3. Perform/Influence SBFS Paradigm.

(Mutually) Converge

This paradigm has the goal of causing each user's emotions to attain a similar, common state.

It can be argued that *Audible Distance* is an example of this paradigm[4]. The simple heartbeat pulse display it employs could be expected to be effective at targeting the unconscious mind; notably, conscious-conscious communication is less important in this paradigm and Audible Distance reflected this by suppressing many of the normal aspects of face-to-face communication. Figure 3-4 shows this paradigm.

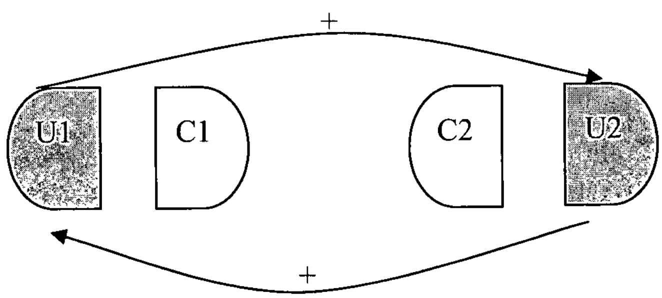


Figure 3-4. Converge SBFS Paradigm.

(Mutually) Understand

This paradigm attempts to bring an understanding of both users' emotional states simultaneously into both users' conscious minds.

The *Galvactivator* fits this paradigm, (especially in the scenario where two people each wear a device) given the stated goals of "facilitat[ing] study of the skin conductivity" and "facilitation of conversation between two people" [8], and considering its simple LED display which is designed to be explicitly comprehended but which would not necessarily have a large direct affective/unconscious impact. Figure 3-5 shows this paradigm.

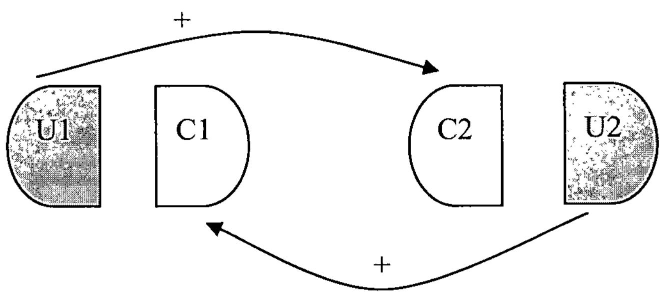


Figure 3-5. Understand SBFS Paradigm.

Control

This paradigm helps user 1 (the controller) learn to exert conscious, intentional influence over the emotional state of user 2 (the subject). Negative/damping "-" arcs U1→C2 and U2→U2 are speculative, and included to indicate that the controller may be assisted if the subject does not know the controller's intentions, and is also unaware of his own unconscious changes; these damping arcs would not be suitable in some contexts. Positive feedback arc U2→U2 is also speculative; this feedback could increase the magnitude of invoked changes in user 2's affective state.

A clear example of this paradigm in the literature has not been identified (although lie detectors have potentially been employed in this fashion), which suggests that an valuable research project would be to build such a system. The system could be subject to interesting tests, with the controller given specific targets for controlling the subject. Fels' theory of *control intimacy* suggests that if a user becomes sufficiently skilled at manipulating a given tool or system, the system will become 'embodied' into the user's identity and controlled unconsciously or instinctively as an extension of the user [17]. This may apply as well to advanced social-emotional control, which would be represented here by a migration of the U2→C1 arc to U2→U1. (A discussion of how connections to the unconscious encourage identification is found in Chapters 4 and 5.) Figure 3-6 shows this paradigm.

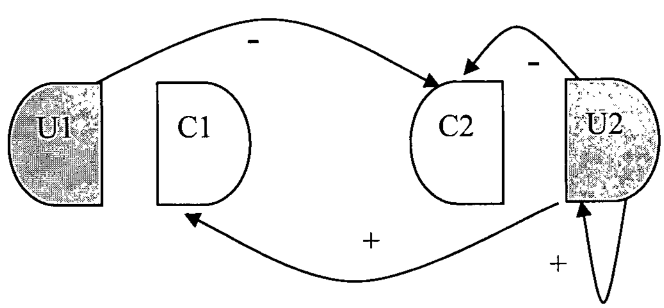


Figure 3-6. Control SBFS Paradigm.

(Mutually) Co-Navigate (the Paradigm of 2Hearts)

This is the paradigm of the 2Hearts system, as discussed in Chapter 9. This paradigm has a novel and more complex goal, which is to facilitate the two users in working together to manipulate their conscious/unconscious state along a mutually negotiated trajectory, with the aim of exploring each other's psychological landscapes as well as creating an engaging experience from the audiovisual display. This paradigm combines aspects of the *converge, control,* and *understand* paradigms, but the diagram is less simplified than the others, as it explicitly shows node SS.

Node SS is important here because one aim of the 2Hearts system is to give users an understanding of inter-user heart phenomena. An assumption of the signal-flow model is that user-user visceral correlation is not usually noticed/processed in natural social interaction. Therefore, arcs to/from SS are not shown as simply augmenting an existing arc. Figure 3-7 shows this paradigm.

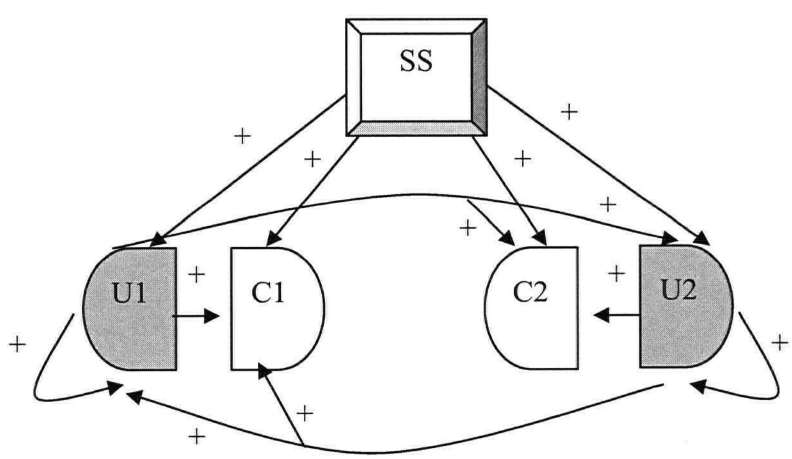


Figure 3-7. Co-Navigate SBFS Paradigm.

4. Signal-Flow Model Development: Emotion and Psychophysiology

Chapters 4, 5, and 6 examine various regions of the SBFS signal-flow model with respect to cited research. This chapter focuses on the relationship between the unconscious (U), conscious (C), and viscera (V) **within** a given user, and discusses implications of this relationship for system design. Chapters 5 and 6 go on to examine between-user links and system-to-user display/perception. As mentioned in Chapter 1, the heart is the main example of visceral function to be discussed here, but the results form a useful foundation for analyzing other types of visceral signal in the future.

4.1 Emotional Processing and the Conscious vs. Unconscious Distinction

Psychological research presents many interesting findings regarding how emotional information is processed by the conscious and unconscious mind. The popular book "The Emotional Brain" by neuroscientist Joseph LeDoux, which presents an overview of such findings, was especially helpful/influential to this aspect of the current project [15]. Relevant findings for the SBFS signal-flow model are presented here.

Initial processing/perception of emotional stimuli (and other stimuli) probably takes place largely in the unconscious.

Cognitive scientists have established that sensory stimuli are processed and filtered by a chain of unconscious mechanisms before reaching awareness [18]. It is now believed that this applies also to the perception and appraisal of emotionally relevant stimuli. Many studies have shown that subliminal or unnoticed stimuli can affect mood and judgement. LeDoux examined the evidence and concluded that:

> "It now seems undeniable that the emotional meanings of stimuli can be processed unconsciously. The emotional unconscious is where much of the emotional action is in the brain". [15:64]

An implication for the signal-flow model is that the path S→Un is justified and important: stimuli from the system do not have to pass through conscious judgement to reach the unconscious. Note that in a sense, all information reaching Cn passes through Un. However, the arcs leading to Cn in the signal-flow model can be retained to stand for signals which are passed through Un relatively quickly and transparently to Cn.

Transmission of emotional information to the conscious depends on indirect and relatively weak routes from the unconscious, including visceral feedback

LeDoux identifies three pathways which transmit emotional information from the unconscious to the conscious: direct (amygdala to cortical) influences, non-specific cortical arousal, and conscious perception of visceral/somatic feedback [15]. He emphasizes the indirect nature of these connections and the possibility for error or misattribution: "[…] the core of an emotion is not an introspectively accessible conscious representation." [15:299] Damioso's 'somatic marker hypothesis' proposes that multi-faceted patterns of somatic and visceral activity are important as feedback leading to conscious emotional experience [19]. The nature of conscious visceral perception remains controversial in the literature; for example Katkin et al. found that subject's ability to consciously estimate their own heart rates was generally poor [16].

One implication for the signal-flow model is that the U→C connection can be considered as relatively weak. Another implication is that the pathway V→C may be of natural importance to emotional experiences, and thus augmenting this pathway with V→S→C might be an effective way to increase conscious understanding of emotion.

> "It now seems undeniable that the emotional meanings of stimuli can be processed unconsciously. The emotional unconscious is where much of the emotional action is in the brain". [15:64]

An implication for the signal-flow model is that the path S→Un is justified and important: stimuli from the system do not have to pass through conscious judgement to reach the unconscious. Note that in a sense, all information reaching Cn passes through Un. However, the arcs leading to Cn in the signal-flow model can be retained to stand for signals which are passed through Un relatively quickly and transparently to Cn.

<u>Transmission of emotional information to the conscious depends on indirect and relatively weak routes from the unconscious, including visceral feedback</u>

LeDoux identifies three pathways which transmit emotional information from the unconscious to the conscious: direct (amygdala to cortical) influences, non-specific cortical arousal, and conscious perception of visceral/somatic feedback [15]. He emphasizes the indirect nature of these connections and the possibility for error or misattribution: "[...] the core of an emotion is not an introspectively accessible conscious representation." [15:299] Damioso's 'somatic marker hypothesis' proposes that multi-faceted patterns of somatic and visceral activity are important as feedback leading to conscious emotional experience [19]. The nature of conscious visceral perception remains controversial in the literature; for example Katkin et al. found that subject's ability to consciously estimate their own heart rates was generally poor [16].

One implication for the signal-flow model is that the U→C connection can be considered as relatively weak. Another implication is that the pathway V→C may be of natural importance to emotional experiences, and thus augmenting this pathway with V→S→C might be an effective way to increase conscious understanding of emotion.

<u>The conscious can be considered as a "serial", symbolic processor with limited bandwidth, and the focus of conscious attention is often steered by the unconscious.</u>

Cognitive load theorists such as Sweller [20] and cognitive scientists generally view the conscious mind as a relatively slow processor that can only attend to a small number of tasks at a given time, whereas the unconscious mind works essentially in parallel [15]. The "spotlight" of attention can be commandeered by the unconscious to focus on emotionally relevant stimuli [15]. The conscious is thought to represent information symbolically, compared to the unconscious's subsymbolic processing [15].

With respect to the signal-flow model, this implies that the total bandwidth of signals reaching C is constrained, and that a 'gate signal' runs from U to C, influencing the selection of signals reaching C. This suggests adding to the representation of C an abstract 'attentional selection' processing module, controlled in part by a gate signal from U. Figure 4-1 shows this arrangement. This perspective will be useful in some of the SBFS design considerations explained in subsequent chapters.

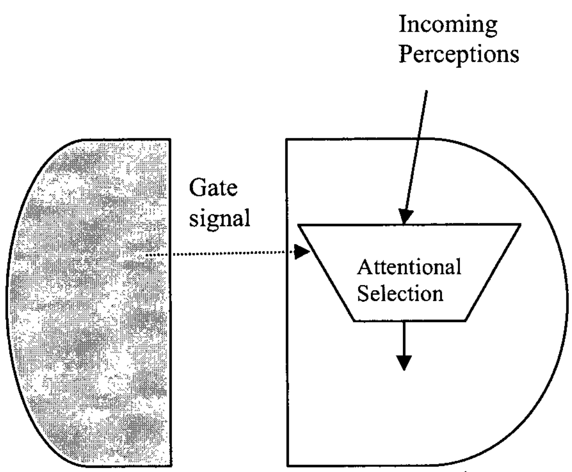


Figure 4-1. Attentional Selection in the Signal-Flow Model.

Conscious representations can lead to emotion stimulation in the unconscious

Conscious representations in short-term and long-term memory can combine to activate the amygdala and other neural emotion systems [15]. This describes the C→U pathway, which in terms of the signal-flow model could be characterized as important but not dominant.

Artificial visceral feedback to the conscious and unconscious can influence and produce emotion; however, non-conscious feedback may be stronger

Studies using artificial audio-based heart rate feedback have shown that false, high heartbeat feedback can bring on panic attacks [21], and that high, false heartbeat feedback in conjunction with visual images can influence feelings of sexual attraction[22]. These studies did not attempt to separate out the feedback paths to the conscious and unconscious. However, studies of subliminal exposure have shown that emotional effects are actually **stronger** when stimuli are presented solely to the unconscious: "our emotions are more easily influenced when we are not aware that the influence is occurring [15:59]"

The implications of this difference between conscious/unconscious feedback are very interesting: in order to convey information to the unconscious, it might be best to prevent this information from reaching the conscious mind. In terms of the model, this information can be used to expand on Figure 4-1 by showing a two-way gate-signal connection between unconscious and conscious mechanisms of perception, as shown in Figure 4-2 below. (In further development of the signal-flow model, the 'gate signal' idea will often be kept in the background as an implied consideration, rather than diagrammed.)

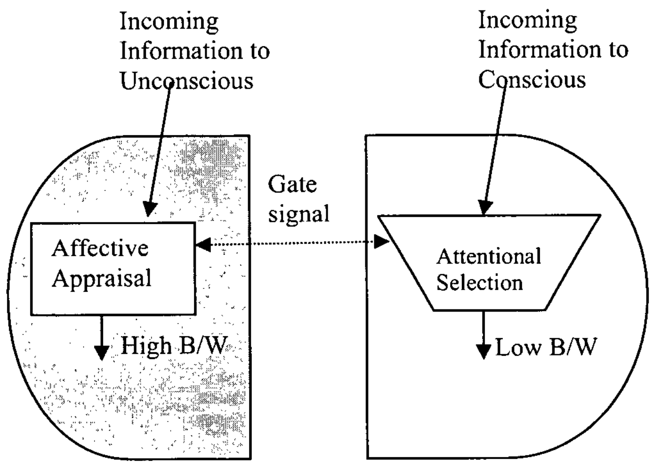


Figure 4-2. The Connection between Conscious and Unconscious Input.

<u>Emotions can be distinguished and described using dimensional or categorical approaches</u>

Many 'theories of emotion' have been proposed which attempt to explain the relationship between different commonly-invoked emotions such as happy/sad/angry/scared, etc.

Dimensional approaches such as Russell's influential "circumplex model" use two or more axes to create a space of emotions [23] [24]. The two most commonly used dimensions which characterize emotional states are arousal (i.e. excited/energetic vs. calm/relaxed) and valence (i.e. pleasant/positive vs. unpleasant/negative). For the purposes of the signal-flow model, a third useful dimensional concept is suggested, grouping concepts from several emotion theories [25] [26] : the concept of *goal-state*. Goal-state is a grouped abstraction of several related concepts (including attraction/repulsion, having/wanting, and expectancy) which could be considered as one or more extra dimensions. Figure 4-3 below shows Russell's circumplex model.

Page 33 has been removed due to copyright restrictions. The information removed was Figure 4-3.

The figure, entitled "Circumplex Model of Affect" shows a 2-dimensional emotion space illustrated by 2 axes arranged vertically and horizontally. The vertical axis is labelled "Activation" (at top) and "Deactivation" (at bottom). The horiztal axis is labelled "Pleasant" (at right) and "Unpleasant" (at left). Emotion descriptor words are arranged circularly around the axes, as follows (listed clockwise from top): alert, excited, elated, happy (upper-right quadrant), contented, serene, relaxed, calm (lower-right quadrant), fatigued, lethargic, depressed, sad (lower-left quadrant), and upset, stressed, nervous, tense (upper-left quandrant).

The figure was reproduced from "The structure of current affect: Controversies and emerging consensus," by L. Feldman-Barrett and J. A. Russell, 1999, *Current Directions in Psychological Science*, Vol. 8, p. 11.

Categorical and prototypical approaches to emotion group various emotional states into 'primary' categories and regard more complex or 'secondary' emotions as subtypes or blends of these categories [25] [27]. These approaches also make use of goal-state concepts to modify or group emotions. Table 4-1 below lists the primary (or 'basic') emotions proposed by a number of researchers.

Page 35 has been removed due to copyright restrictions. The information removed was Table 4-1.

The table, entitled "A Selection of Lists of 'Basic' Emotions", contained 3 columns: "Reference", "Fundamental emotion", and "Basis for inclusion". Each row lists a certain researcher(s), the basic emotion description-words proposed by that researcher (such as "Anger", "disgust", "fear" "joy" etc.), and the researcher's stated rationale (such as "Hardwired", "Unlearned emotional states", etc.)

The table was reproduced from "What's Basic About Basic Emotions?", by A. Ortony and T. J. Turner, 1990, *Psychological Review,* Vol. 97, No. 3, p. 316.

It is suggested that these two dominant approaches to emotion theory (dimensional and categorical) can be employed to add detail to the signal-flow model. In particular, the categorical description seems more aligned with the 'symbolic' representation used by the conscious mind, while the dimensional approach might relate more strongly to the U←→V connection.

This perspective can be illustrated by adding internal detail to U and C, showing an 'affective dimensional variables' module in U which communicates several dimensions of information with V, and an 'affective categorization module' in C which creates a conscious/symbolic 'affective representation'. The affective dimensional variables are viewed as continuous and dynamically changing, whereas the affective representation is viewed as changing more slowly between discrete states.

Other findings discussed above can be incorporated here by showing the (relatively weaker) links from conscious representation back to unconscious affective state, from visceral action to conscious awareness and from unconscious state to conscious awareness. Figure 4-4 depicts these details added to the signal-flow model.

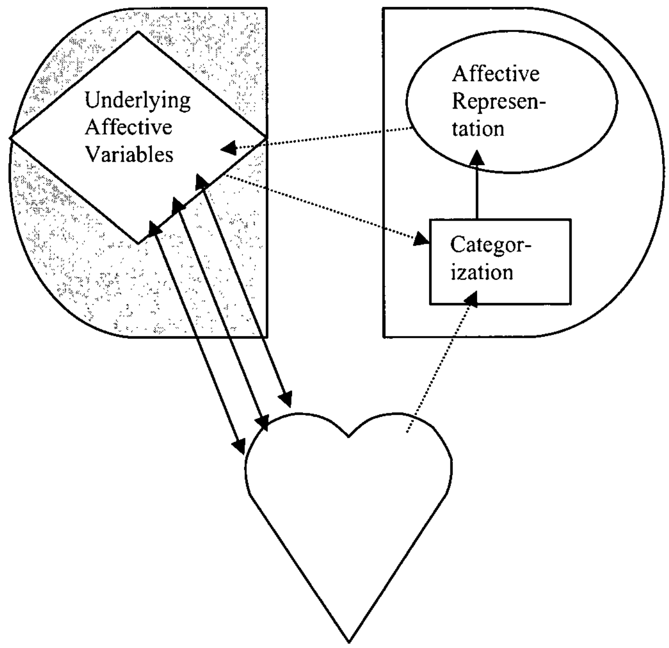


Figure 4-4. Conscious vs. Unconscious Representations of Affect.

The assumption of underlying, dynamically varying 'affective variables' in the unconscious is an important and somewhat speculative aspect of the SBFS signal-flow model. It will be shown throughout the remaining chapters that this characterization usefully links together many SBFS theoretical and design issues.

Employing both a dimensional and a categorical approach to emotion in the same model can be seen as an application and extension of Picard's advice when she states:

> "Moreover, the question of whether to try to represent emotions with discrete categories of continuous dimensions can be considered a choice, as each representation has advantages in different applications" [1:169].

With respect to Picard's statement, it is proposed that SBFS's are an application that benefits simultaneously from both representations.

4.2 Psychophysiology and the Heart

This section uses findings from psychophysiology to examine how psychological influences cause measurable effects on the action of the heart. These findings are relevant to arcs U→V and C→V, and especially relevant to designing the system input pathway V→S.

Biology of Heart Control

According to Hugdahl [16] the connection between the heart and the unconscious/emotional mind stems from the connection between the heart and the autonomic nervous system (ANS). The autonomic nervous system functions primarily without conscious control and thus can be considered part of unconscious node U in the signal-flow model. The ANS has two branches: the sympathetic system, involved with arousal, and the parasympathetic system, involved with relaxation. The sinoatrial (SA) node of the heart is enervated by both sympathetic and parasympathetic nerves, which carry signals to speed up and slow down heart rate, respectively. Heartbeats are created by a wave of electrical depolarization initiated in the SA, which sweeps diagonally across the heart.

Measurement of Heartbeats

Many psychologically-linked aspects of heart behaviour involve changes in the beat-to-beat timing of heartbeats, although the force of each heartbeat and the contraction vs.

time profile of each individual beat are also of psychological interest. The present research has focused on time-of-beat related measures, due to the availability of cheap and accurate sensors to detect time-of-beat (see Chapter 8), and the satisfactory number of different psychophysiological phenomena that can be explored using time-of-beat. Time-of-beat can be measured with electrical sensors such as the electrocardiogram (ECG), which employs electrodes on the chest to measure the electrical activity of each heartbeat. Each beat measurement typically includes an identifiable spike known as the 'R' peak, and the common measure Instantaneous Heart Rate (IHR sometimes referred to generally as HR) is given from the inverse of the interbeat intervals (IBI), measured from R peak - R peak [16]. Heart rate exhibits both tonic and phasic psychological effects [16].

Heart Rate Variability (HRV) is a general term for several different oscillation/variation phenomena in the HR, also linked to the action of the sympathetic/parasympathetic nervous systems [28]. Simple standard measures of HRV include Standard Deviation in IBI and Standard Deviation in Delta-IBI [29]. The frequency spectrogram of HR can also be used to derive measures of HRV such as high-frequency power, low-frequency power, and high/low power ratio [28]. Some newer and more experimental heartbeat measures are based on entropy and fractal mathematical techniques.

Psychophysiological Heart Rate Phenomena

Long-term tonic HR is related to many factors, including gender, age, fitness level, and daily biological cycles. Phasic elevation of HR is caused by excitement/arousal and by stressor events [30]. Known stressor events include performing mental arithmetic, experiencing cognitive dissonance, being exposed to physical pain, and performing challenging motor tasks [16]. Certain 'hyperreactive' individuals experience larger phasic HR elevation in such situations- this tendency is associated with a certain personality profile and with greater risk of heart disease [16]. This finding indicates that different SBFS users might have different signal strengths along arc V→S, perhaps requiring compensation or calibration by the system.

Several other event-triggered patterns of acceleration/deacceleration have been observed in the HR. An 'orienting response' occurs after an unexpected stimulus: HR first goes down below baseline for roughly 2 s (thought to indicate the focusing of attention), then returns to baseline or higher (thought to be related to emotional/internal processing) [16]. Anticipation of an expected event also causes a distinctive deacceleration [31]. Heart rate has been found to increase with 'appetitive motivation' compared to skin conductance which increases with 'aversive motivation' [32]. Attitude with respect to 'sensory intake' (or 'attentional stance') is thought to be reflected in HR patterns: an 'open/externally oriented' state lowers HR while a 'closed/internally oriented' state raises HR [33] [34].

Comparing the phasic-HR-related findings to the models of emotion discussed in Section 4.1, it is clear that HR responses relate to the emotional dimension of arousal and also to goal-state (anticipation, motivation etc.). A connection with emotional valence seems plausible but less obvious. Interestingly, though, the ideas of 'orientation' and 'attentional stance' seems to also be connected to the balance in activity between nodes U and C, and the state of the proposed 'input gate'. In other words, signals from the V nodes may give information about the combined state of U, C and attentional orientation.

Regarding heart rate variability, the HRV literature is extensive, complex, and containing of some controversial topics, but a general summary might be that high HRV is associated with low arousal and positive affective valence. High resting HRV is thought to indicate a general disposition towards parasympathetic activity [35]. It should be noted that one of the strongest frequency components in HRV is synchronized with breathing (this is known as respiratory sinus arrhythmia [16]). This connection means that voluntary breath control can affect HR/HRV; and breathing patterns during speech/socialization might affect HR/HRV.

Statistical/Machine-Learning Techniques Applied to Physiological Variables

A growing number of researchers are employing statistical techniques in an attempt to match combinations of physiological measurements with emotional states (usually those emotional states defined by categorical theories), with encouraging success. Ekman and Levenson showed that patterns of autonomic nervous system activity can distinguish among emotions [36]. Collet used combinations of EDR, temperature and respiration to discriminate successfully pair-wise between emotions [37]. Picard used a hybrid method of fisher projection and sequential forward floating selection on EDR, respiration and ECG to distinguish happiness, sadness, anger, fear, disgust, surprise, love and neutrality with 81% overall accuracy [38].

Psychophysiology in the Signal-Flow Model

Findings from the above psychophysiology research can be incorporated into the signal-flow model to add detail to the U←→V connection. Taking heartbeat as the example of visceral function, the underlying affective variables are shown as influencing the action of the ANS. The ANS connects to a particular organ through a dimensionally-reduced set of variables (for instance, sympathetic activation and parasympathetic activation). The ANS influence on the organ is then modulated by the electrical and physical characteristics of the organ to produce the particular signals which can be directly measured with sensors. Figure 4-5 shows this arrangement.

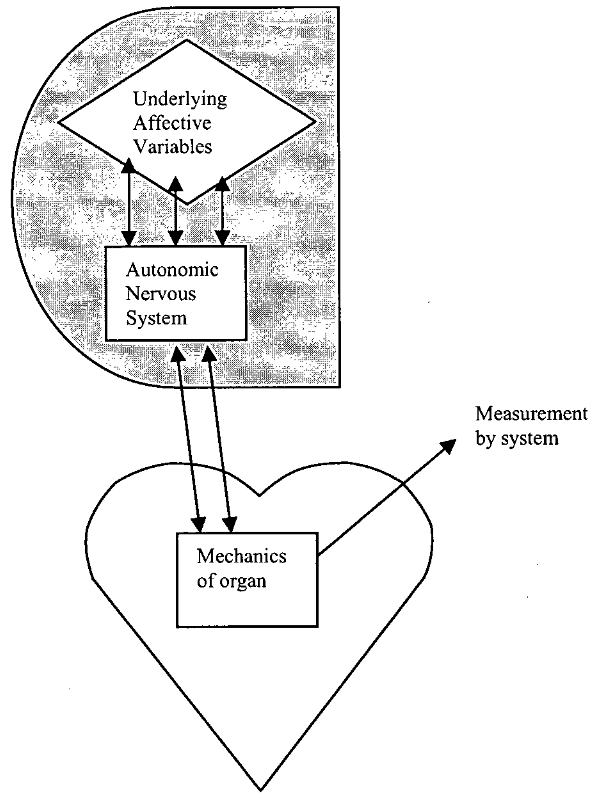


Figure 4-5. Different Levels of Unconscious Representation of Affect.

4.3 Relevance of Natural/Internal Feedback Pathways to SBFS Input Design

Having examined the different signals and pathways that connect emotion, conscious experience, and visceral action, the implications for design of system input channels Vn→S can be considered.

A major theme of the research discussed is that conscious/unconscious feedback from visceral action can have significant effects on unconscious emotional state, and on conscious emotional appraisal; furthermore, it appears that artificial (real or false) feedback using display modalities such as audio can engage similar brain mechanisms as the mechanisms which normally interpret natural/internal feedback.

This is very encouraging, suggesting that SBFS feedback through pathways V→S→U and V→S→C can convey information effectively to the conscious and unconscious mind, as desired for effective system operation in the various SBFS paradigms. It remains in further chapters to discuss output/display mechanisms for the SBFS; however, the general principle is established: *the pathways V1→S→Un and V2→S→Cn will be effective if they can simulate or mimic the action of natural pathways V→U and V→C, respectively.* (Note that this finding is generalized to include both intra-user and inter-user feedback. This is discussed further in Section 5.1.)

One step in mimicking such pathways is to derive from sensor data a set of variables which are similar to the variables taken as input by the various conscious/unconscious mechanisms of emotional appraisal. As, Figures 4-3 and 4-4 above show, such variables of interest include the conscious affective representation/state, the variables describing the physical organ action, the variables describing the ANS action, and the underlying affective variables. Thus, one useful processing goal for SBFS is to approximate the categorical, conscious affective representation, while a second distinct goal is to undo or 'invert' the various signal transformations occurring between the unconscious/underlying affective variables and the visceral measurements.

5. Signal-Flow Model Development: Inter-User Pathways and Social Environment

This section provides a more detailed analysis of signal-flow pathways that may exist **between** the users in the context of a SBFS. Some of these pathways describe normal, social, verbal/non-verbal communication, which has been researched in several fields such as psychology, sociology and computer-supported collaborative work. Some of these pathways only occur in the presence of a SBFS; thus, speculation/theory in those areas is more novel. Such novel areas include:

- Impact/interference of SBFS physical design on the social context of interacting users.
- Implications for SBFS's of user intentions toward each other
- Heartbeat phenomena detectable by jointly considering the heartbeat signals of two interacting users.

5.1 Pathways of Non-Verbal Communication

As mentioned in Chapter 3, in the signal-flow model, arcs proceeding from En represent communication signals from user to user, both explicit and implicit. This duality of function of the En nodes was chosen because the behaviour of En consists of bodily modulations which for the most part can be controlled and perceived **either** consciously or unconsciously. For instance, tone of voice can be used purposefully for dramatic or comic effect; or it can unconsciously betray someone's true emotions even against their will. Likewise, someone may feel cheered due to someone else's happy facial expression, even if the perceiver is not consciously attending to that aspect of the conversation.

Note that under the signal-flow model, the unconscious modulation of somatic muscles is represented by either Un→Vn or Un→En depending only on *whether the modulation is externally perceptible*: thus there is much overlap and similarity between these two pathways. Therefore, many of the previous findings regarding somatic modulation by the unconscious apply similarly to perceptible somatic modulation.

Given the previous discussion of close links between emotion and the unconscious, it is not surprising that typical social communication of emotion is thought to occur along the pathway U1→E1→U2. This unconscious, physically expressed communication is sometimes called *sentic modulation.* As stated by Picard,

> "Sentic modulation [...] is the primary means of communicating human emotion. In fact, few people are good at articulating their emotional state, but expressing it through sentic modulation is natural, and usually subconscious." [1:25]

Picard later states, "information presented through the 'affective channel' does not usually demand conscious attention". [1:56]

A useful specific example of this implicit affective channel is the communication of affect through facial expression. The facial muscles can be controlled either voluntarily or involuntarily [39], and Ekman argues that the involuntary (i.e. unconscious) movements are those that truly express emotion [40]. Various studies of brain-damaged patients suggest that separate nural pathways are involved in conscious/unconscious facial expressions- the frontal cortex is implicated in voluntary control while the basal ganglia is implicated in unconscious expression; although there is some overlap [41]. Likewise, studies suggest that the emotion content of facial expressions can be processed unconsciously [42] [43], and that such unconscious emotional processing may engage different neural pathways from expressions that are consciously analyzed [44].

Thus, C1→E1→C2 could be considered the 'explicit/verbal pathway' and U1→E1→U2 the 'implicit/affective pathway'. (See Figure 5-1.) (The existence and relevance of U1→E1→C2 is not discounted, however- a description of this pathway has been identified in research regarding *cognitive empathy,* to be explored below.)

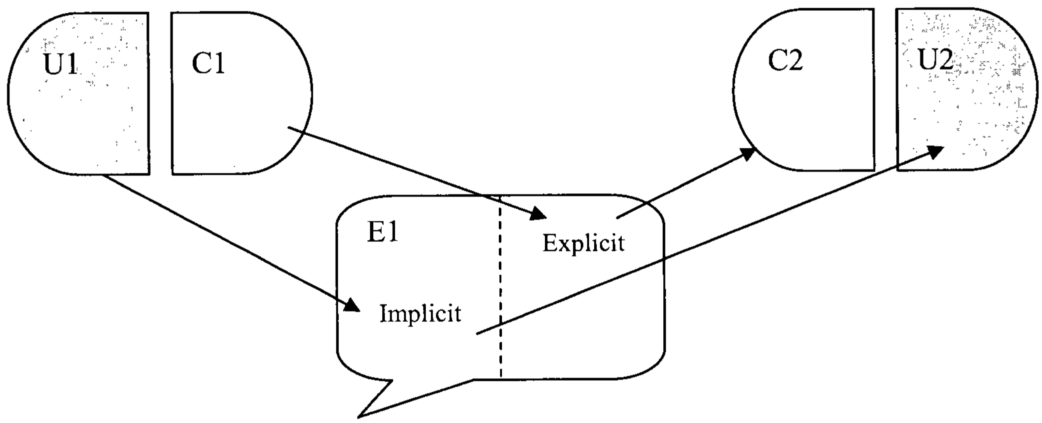


Figure 5-1: Implicit/Explicit Pathways via the External Communication Node.

The Essentic Forms Theory of Sentic Modulation

The communication of verbal/explicit information is fairly well-understood, so it is not examined further in the present work. However, it would be useful to have a theory detailing the type of signal that is transmitted along the pathways U1→E1→U2 and U1→E1→C2.

The researcher Clynes developed an interesting theory regarding how emotion is expressed non-verbally, claiming that characteristic 'essentic forms' are used by humans to express given emotions, and the forms can be recognized across a variety of display formats, for example: curves and shapes in paintings, position vs. time contours of hand movements, and rhythms in music [45]. A device called a *sentograph* was used to measure patterns of finger pressure produced by subjects as expressions of different emotions.

The work has been controversial throughout the years, but notably, Picard discusses/defends Clynes' research and considers that it is relevant to Affective Computing [1]. The idea of essentic forms, if valid, is attractive to employ in the signal-flow model since it involves low-dimensional, time-varying signals that can be transformed/displayed in many ways but still convey affective content. A suggested characterization of essentic forms with respect to the signal-flow model is that these forms are derived within the unconscious from the underlying affective variables, and transmitted to external physical modulation.

Given that essentic forms are characterized as single-dimensional time-varying quantities, it is speculated that the mechanism which converts underlying variables to essentic forms acts essentially to convey the most dominant or notable time-varying contours of the trajectory of underlying variables. Therefore the term 'essentic envelope transform', used in an abstract/metaphorical sense, is proposed for this mechanism.

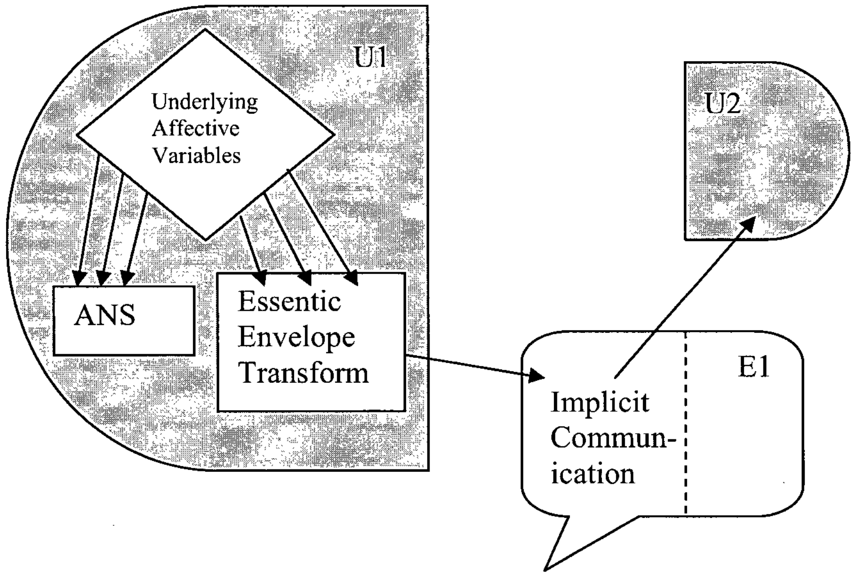


Figure 5-2. Implicit Communication Through the Transmission of Essentic Forms.

It is likely that this essentic envelope transform module transmits similar information to externally-visible somatic signals (node E) as well as hidden-somatic signals (part of node V). This suggests that an SBFS taking hidden-somatic signals as input would be dealing with data very similar to the 'essentic forms'; perhaps such data could be easily communicated with little extra processing. An open research question is whether the essentic forms are tied to modulation patterns measurable from nervous system/brain action. Neuropsychological research involving this question could be of great value to SBFS's and related systems.

One caveat regarding the essentic forms theory is that it is not clear whether these forms are more associated with conscious production (U1→C1→E1), unconscious production (U1→E1), or both [1].

Empathy, Modeling and Identification

Recent psychological research has defined two kinds of empathy: *affective empathy*, 'experiencing an emotion as the result of someone else's mental state', and *cognitive empathy*, 'understanding and predicting someone else's mental state' [46]. It is apparent that these concepts match well with the signal-flow model descriptions of U1→E1→U2 (affective empathy, thus identified with the affective pathway) and U1→E1→C2 (cognitive empathy).

An interesting aspect of cognitive empathy is that it involves a process of psychological modeling; one person builds up a conscious model of the other's emotional state. Compared to affective empathy, cognitive empathy involves a greater degree of separation from the object: the 'cognitively empathizing' person understands the perceived emotion-model as being separate from her own emotions. On the other hand, affective empathy involves a kind of identification with the object: the 'affective empathizer' does not necessarily recognize whether the source of an emotion comes from the self or the other. The term *emotion contagion* has been used to describe affective empathy without conscious recognition of the process [47].

In Chapter 4 it was shown how consciously experienced emotions are achieved through a mechanism of categorization and symbolic representation. The signal-flow model can be extended to accommodate the social phenomena of modeling and empathy by introducing multiple symbolic representations in the conscious mind: a representation of the self's affective state, and a representation of the other's affective state. Figure 5-3 below shows these multiple representations. The unconscious is not viewed as maintaining such a discrete separation between affect originating from self or from other, thus information that flows into U may or may not lead to correct conscious attribution via pathway U→C. This theme of unconscious identification vs. conscious modeling will re-occur in the study of music perception in the next chapter.

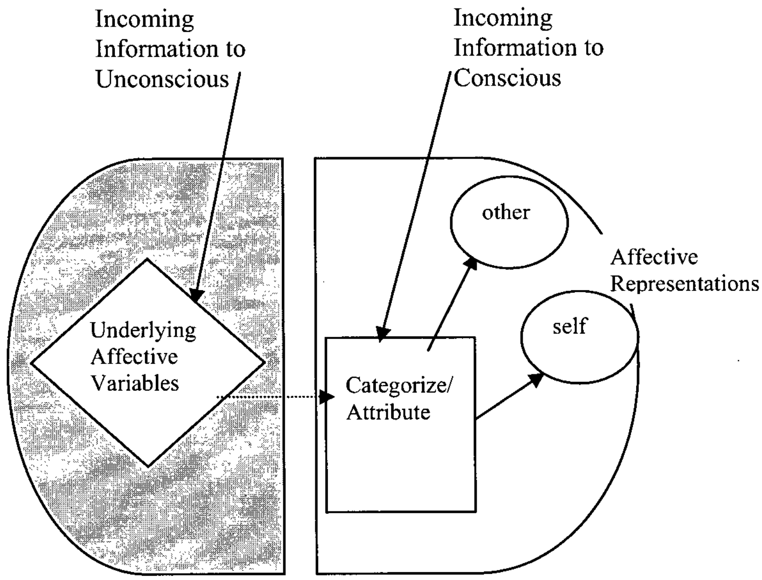


Figure 5-3. Identification vs. Modeling in the Signal-Flow Model. Information directly reaching the unconscious is less likely to be attributed to the affective state of the 'other'.

Non-verbal communication as a model for SBFS display channels

Since it represents a natural communication of emotional information, the implicit pathway could be useful as a model for the artificial SBFS output pathways S→Un. However, it was also suggested in Section 4.3 that SBFS system output could be inspired by natural intra-user feedback paths. Thus an interesting decision arises- should the system attempt to mimic internal visceral feedback, or should it mimic nonverbal communication? In fact, there is no dilemma, because elements of both schemes can be combined (and furthermore, as implied above, nonverbal communication already mimics visceral feedback to some extent).

An advantage of mimicking non-verbal communication is that non-verbal communication parameters (such as pitch inflection in voice) are more easily measured, and thus copied. On the other hand, for a face-to-face SBFS such as 2Hearts, a disadvantage of mimicking non-verbal communication is that the system display may interfere with natural communication (because the two channels will be competing for the same perception/processing resources in the brain). A further point is that non-verbal communication itself mimics within-brain affective communication to an extent.

The approach that was taken in 2Hearts design was to draw inspiration from both natural models, as well as a third form of emotional communication that is 'artificial' but commonly used- **music.** This strategy is discussed further in the next chapter. In general, the SBFS designer is free to draw upon all three areas of research.

5.2 Interference/Environment Context

Because the SBFS will have physical components such as sensors, digital processors and display devices, which may be located in proximity to the users, and because the system may emit audio/visual signals that interact or interfere with the communication signals between users, the system should be considered to have a potential direct effect on the 'natural' social arcs E1→C2 and E1→U2. This can be illustrated in the signal-flow model by showing the system's boundary as an expanded outline, intersected by the inter-user arcs, as in Figure 5-4.

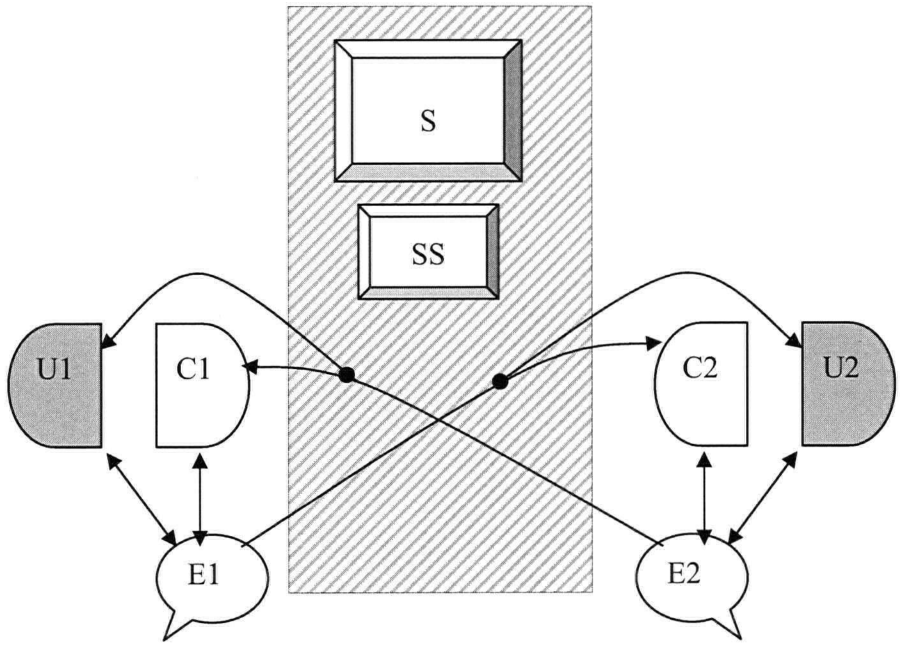


Figure 5-4. Impact of SBFS on Social Environment/Context. The system influences the interaction environment, through audio/visual interference, constraint of users' movement or vision, or the creation of ambience/ emotional connotations.

In some cases, interference with natural social communication pathways can help a SBFS meet its design goals. In *Audible Distance,* the system visual displays also served to block normal visual perception, thus damping arcs E1→C2 and E1→U2 [4]. For SBFS designed to efficiently support certain tasks, some interference may also be acceptable, since computer-supported collaborative work research has shown that task-efficiency does not necessarily depend on all communication channels being present [48].

However, for systems where the SBFS is intended to augment rather than replace normal, face-to-face social interaction, such interference is undesirable. Therefore effort must often be made to design SBFS with unobtrusive, un-interfering input and output technology. This group of concerns will be discussed further regarding the design of the 2Hearts system, in Chapters 8-9.

5.3 Intentions/Expectations/Privacy

Not yet discussed is the role played by users' general *intention* in using the SBFS. The variety of intentions are similar to the paradigm goals, although they may shift moment to moment or be different from what the system designer had in mind. For instance one user can intend to *control* the other user or instead to *understand* them. Likewise one user may know the other user's intentions and seek to *resist* them or *co-operate* with them. These shifting conscious or unconscious attitudes or stances may affect the strength of both input and output pathways to the user, via shifts in focus of attention or modulation of expressive signals. Such effects of inter-user attitude are a topic for further study.

A related topic is one of privacy and willingness to use a SBFS's; for instance, one objection might be that people have evolved to only display a limited amount of visceral/ unconscious information for good reasons, and that they may value this form of privacy when engaging in social interaction. Picard points out this concern (regarding biometric wearables) and states that it is important for the user to "[retain] ultimate control over the devices he chooses to wear, so that they are tools of helpful empowerment and not of harmful subjugation" [1:244]. This important topic warrants further explanation; at this early stage of exploring SBFS designs, it can be addressed by gathering and taking careful note of test users' observations and opinions regarding privacy, comfort, embarrassment, etc.

5.4 Hypothetical Interaction Phenomena Between Heart Signals of Socially Interacting People

As shown in Section 4.2, there has been considerable research probing how measures of a single person's heart beat signal respond to various stimuli and conditions. However, there has been less study into phenomena that can be detected by analyzing the heartbeat signals of two people interacting. It can be hypothesized that for two people engaged in social interaction, their biosignals will sometimes vary with similar patterns, due to the effects of empathy leading to common emotional states. Such phenomena could then be processed by SBFS system node SS, and conveyed to the users.

A promising recent result for SBFS systems is a study by Marci and Orr, who found that in therapy sessions where empathy between therapist and patient was rated as higher, the 'concordance' (based on global and episodic Pearson's correlation calculations) between patient-therapist EDR signals was greater [49]. Regions of mirrored EDR behaviour between users has also been observed in Conductive Chat data [9].

Regarding the heartbeat signal, several types of specific possible synchronization and correlation are listed here, each of which merits further study:

Heart Rate Co-Varying Behaviour: Loosely defined, a tendency for heart rate 1 to be higher when heart rate 2 is higher. This type of phenomenon is strongly predicted to occur for the simple reason that elements of the environment or the interaction that cause arousal in user 1 are likely to cause arousal at a similar time in user 2 (for example, something humorous or offensive is said, or a loud noise occurs in the environment).

A simple measure of co-varying heart rate is the zero-lag cross-correlation, (or Pearson's correlation) which will indicate whether two heart rate signals tend to deviate from their respective means in the same direction at the same time. This measure varies depending on the time window employed: for example, if two users both experienced remarkably high but steady heart rates during a certain episode, a correlation window shorter than the episode would fail to show high correlation. On the other hand, long correlation windows may emphasize longer-term trends over short-term episodes.

An interesting way to work around this limitation might be to calculate correlation based on heart rates which are pre-normalized relative to long-term observations of per-user heart rate behaviour; then a shorter correlation window could be used for dynamic observations. This technique may be investigated in the future based on heart rate data that has been collected in the current work.

Examining the cross-correlation at lag offsets other than zero can give information about delayed-reaction effects where an emotional behaviour by user 1 causes a delayed emotional response in user 2.

An alternative measure of heart rate co-varying behaviour is *coherence,* a measure which studies synchronization behaviour at different frequencies. Coherence measures warrant further investigation for SBFS application, especially since limited-bandwidth HRV phenomena are important to heart psychophysiology (see Section 4.2).

Simpler instantaneous measures derived from heart rate can also portray common heart rate behaviour: the difference between users' heart rate roughly indicates the amount of momentary disjunction between the users' emotional states (if heart rates are normalized per-user, the information is more useful), and the mean of two users' heart rates might be used to indicate the overall importance or interest level of a given moment of social interaction.

Phase Synchronization: Also referred to as entrainment. It occurs if beats from the two users fall close together in time more often than would be expected by chance. This has been observed experimentally by van Leeuwen et. al to occur between mothers and fetuses [50] Entrainment has also been shown to occur between a person's heart beat and an external audio or visual driving signal [51] The importance of entrainment to emotion or social interaction is not currently known. Phase Synchronization can occur at similar absolute frequencies, giving a 1:1 heartbeat ratio, or at dissimilar frequencies, given a m:n ratio between the heartbeats, where m and n are integers. Phase synchronization has not yet been quantitatively explored in the current work; however a 2Hearts system design consideration has been to ensure that episodes of phase synchronization are perceptible in system display should they occur (see Chapter 10).

Other Synchronization Measures: The measures *Cross-entropy* and *mutual information* have also been identified in the literature as potentially useful measures of heart-heart synchronization, and merit further study for application in SBFS design/testing [52] [53].

6. Signal-Flow Model Development: Audio Perception and Display Pathways

This chapter develops the signal-flow model by analyzing the pathways connecting system output to user input, focusing on the output modality of musical audio. Research relating to audio/music perception, auditory display, and emotional/physiological effects of music listening is examined. In this way, the cyclical analysis of the 'user-to-system-to-user' pathways is completed.

For the designer of a SBFS system, the most extensive aspect of the signal-flow model under the designer's control is the display connection from system to user; therefore a detailed sub-channel analysis is developed for these display pathways.

Regarding the distinction between 'audio display' and 'musical display', the definition of music is an unsettled topic of philosophical debate. Here it will simply be assumed that any audio display representing structured patterns of physiological activity can be considered music in a sense; and that findings from musical research are relevant to the goal of conveying information/meaning through sound.

6.1 Audio/Music Perception and Auditory Display

A SBFS can synthesize or manipulate various parameters of sound and music in order to convey information to the listener; the use of audio in the display of data to human perceivers is known as *auditory display [54]*. A list of available parameters to consider can be constructed by considering human perception of sound, beginning at the lowest level.

Biological Basis of Audio Perception

Sound exists physically as longitudinal pressure waves travelling through the air or another medium. Air pressure waves are converted into mechanical vibrations by the organs of the ear, causing the vibration of specialized hair cells within the cochlea [55]. Each hair cell resonates within a certain narrow frequency range due to its placement along the basilar membrane of the cochlea, and is connected to a nerve cell which sends more or fewer nerve pulses along the auditory nerve depending on the amplitude of sound within that frequency range. It has been shown that neurons in the auditory cortex of the brain, which react to stimulation from the auditory nerve, are physically distributed according to a progression of frequency ranges (this is called a *tonotopic* arrangement) [56].

Thus, the fundamental parameters of sound at the lowest level of human perception are time and frequency distribution. However, brain processes very early in the chain of processing act to derive other parameters, such as spatial location, volume, pitch, and brightness.

Lower-Level Audio Display Parameters

This section briefly describes various forms of audio manipulation and some related findings regarding human perceptual capabilities. The continued collection and summary of perceptual-ability measurements regarding these variables would be valuable future work.

Volume

Volume perception is related to the amplitude of an audio waveform, but perception varies for different frequency ranges (this variation is described by the *equal-loudness curves*) [57].

Accuracy for volume discrimination as measured by just-noticeable difference (JND) is about 1dB, while dynamic range varies from roughly 45 dB at 30 Hz to105 dB at 3,000Hz. [58] It has been observed by the author that moment-to-moment volume changes (including oscillating volume, known as tremolo) are generally noticeable, but the ability to compare or remember exact volume levels at different times is less good. Quantitative findings regarding long-term memory for volume would be useful to identify.

Pitch

The perception of fundamental pitch is related to frequency distribution, but there is not a simple mapping between the two. The pitch perception system looks for distinctive overtone patterns like those produced by vibrating physical objects [59].

Musical scales of pitch are related logarithmically to frequency, stemming from the fact that two frequencies are heard as more related or 'consonant' if there is a simple mathematical ratio between them. The unit of musical pitch is the *semitone*, which is divided into 100 *cents*. [60] Typically, musical scales include a classification that repeats with every frequency doubling (or octave). Pitch can be perceived simultaneously in two different ways: *pitch height* refers to frequency in absolute terms while *pitch chroma* refers to pitch relative to the repeating musical scale: in other words, tones of 440Hz and 880Hz have the same chroma.[61]

Pitch has good long-term memorability: one study involving learning a test pitch and recalling it a week later found a typical error of +/- 1 semitone [62]. The JND for pitch perception is around 10 cents, i.e. a deviation of approximately 0.6% [58]. However pitch perception becomes less precise toward the low and high extremes of perceptible frequency [61]. The dynamic range of pitch is around 20kHz [58].

Brightness

Brightness describes the balance of high/low frequencies in the spectrum of a given tone. As in the case of volume, it has been observed by the author that brightness modulation can be perceived, but longer-term retention/comparison of brightness values is less good. Brightness often has affective connotations: bright sounds sound 'harsh' or 'tense', bright speech sounds 'less intelligent'[63]. Brightness could be considered one element of timbre, discussed next.

Timbre

Timbre is often defined by exclusion as 'all parameters except the others listed' [63]. Timbre essentially describes the 'flavour' given by a particular shape of the frequency spectrum. Timbre signature helps distinguish different sources of audio production within the total perceived audio signal. Grey showed that a useful characterization of timbre perception can be given in terms of clusters in a three-dimensional "timbre space" [64].

Mathematically generated tones often have frequency spectra which are brighter than natural sounds and lacking in subtle, chaotic variation over time. If this is not ameriolated through careful sound design, tones can sound harsh and unnatural.

Spatialization

The spatial origin of a sound is estimated within the auditory cortex according to several factors including the time-lag between sound arriving at the left and right ears, the amplitude difference between the ears, and the angle-dependant frequency filtering effects of the head and the pinna (exterior structure of the ear) [58]. At a slightly higher level of processing, environmental interactions such as echoes (causing reverb and delay) also contribute to perceived sound source location [58].

Envelope

Typically musical audio is perceived as segmented into discrete entities, delineated by amplitude envelopes which begin and end at low or zero amplitude. Other parameters such as pitch and timbre may also vary through the duration of a segment with certain envelope shapes. An envelope shape is often described in terms of attack time, decay time, sustain level and release time. Envelope interacts with timbre to determine a note's perceived 'flavour' or 'character'.

Timing

When significantly perceptible volume envelopes occur, onset times for each discrete note are perceived. Two notes separated by about 20ms or more can be perceived with the correct time-order [58].

Higher-Level Sound Parameters

Human perception has a strong ability to identify and track audio patterns created by the manipulation of the above variables through time. These high-level patterns contribute to an affective/aesthetic experience, even without formal musical training on the user's part or a conscious/explicit understanding of what the patterns are.

Rhythm

Rhythm is the perception of patterns in timing information, defined by Bargar as "a sequence of durations measured in micro and macro units" [65:162]. Rhythms can be very memorable, and two rhythm patterns related in a mathematical way can often be perceived instinctively as variations or transformations of each other.

Tempo

Tempo describes a perceived sense of 'fundamental frequency' in rhythm, although this is hard to define algorithmically. The JND for tempo has been measured at 6.2%-8.8% [62]. The long-term tempo memory of the average person was shown to be very accurate for remembering favourite songs, with 72% of subject's estimates falling within 8% of correct tempo. It is interesting to note the similarity of typical musical tempo (50-150 bpm) to typical heartbeat frequency (50-150bpm) and their similar roles as providing an 'underlying fundamental rhythm'; this was also pointed out by McLaughlin [66]. This apparent connection is relevant to investigate in SBFS design.

Harmony

Harmony involves the interaction of pitches with other pitches. A combination of notes will be perceived as more or less consonant (which tends to sound pleasant and stable) or dissonant (which tends to sound unpleasant, tense, or unstable). Consonance/dissonance depend on the fractional relationships between pitch frequencies and how those relate to the natural overtone series [67]. Harmony can occur between pitches in parallel, in series, or even between pitches that are not present but 'implied'. Cultural/learned harmonic rules and contexts can be developed, and influence the perception of pitches.

Musical Structure

Music is formed by repeated patterns of the above parameters, which undergo various structural transformations, and are sometimes unexpectedly changed/interrupted, as a piece progresses . Meyer's influential work emphasized the role of expectancy maintenance/violation in the perception of music [68].

Pitch and Time as Morphophoric Mediums

Shepard points out that the dimensions of pitch and time can be described as *morphophoric mediums,* meaning mediums "capable of bearing forms" [61:154]. In other words, structures in pitch and time can be transposed and modified in various ways, and these structures will be perceived as having related identities, with the nature of the modification appreciated in some way. Pitch in particular supports a complex variety of meaningful structural transformations, due to the complex set of consonant/dissonant and scale relationships between the pitches of various notes.

Other dimensions such as volume and timbre support recognizable structural transposition but to a lesser extent; as Shepard states, "we have melodies in pitch, but no real analog to melody in loudness. [...] The same degree of structural information is not preserved in loudness as it is in pitch and time." [61:155]

Perceptual Grouping and Separation of Sound Streams

Audio is usually perceived as being divided into 'streams', which group certain frequencies and sound parameters at a given time as belonging to a common source. This capability may be related to survival-oriented 'real-world' sound perception, in which it would be beneficial to be able to track and attribute multiple sources of sound throughout time.

The perceptual grouping of audio data into a common stream is promoted by: common timbral attributes, common spatialization attributes, common participation in a given element of musical structure, and gestalt principles such as symmetry, good continuation, and common fate [63] [69] [70].

Other Factors in Auditory Display

Time Scale Parameter Nesting

Research has shown that an effective technique for increasing the number of audio parameters for system display is parameter 'nesting': modulating the same parameter at different time scales [63]. For example, slow variations in fundamental pitch can indicate one variable while the amplitude of faster oscillations of the pitch ('vibrato') can indicate another variable.

Separation of Frequency Ranges

A well-known technique in composition and arrangement is to utilize instruments with distinct inherent frequency ranges to play contrasting/interacting melodic parts. This enhances the ability to simultaneously appreciate each separate part. Thus, multiple streams of audio information may be perceived with greater accuracy if the streams occur in distinct, separated frequency ranges.

Audio Display Compared to Other Modalities

Kramer has compiled useful lists regarding the advantages and disadvantages of employing auditory display compared to other modalities [54]. Many of the factors identified are of obvious relevance to the use of audio in SBFS design. These points are summarized here with brief explanations adapted by the present author. Note that many of these points relate to previously discussed phenomena.

<u>Advantages of Auditory Display for SBFS use [54:9]:</u>

Eyes-free: Auditory display does not require visual attention, thus normal visual contact between users can be maintained.

Backgrounding: Conscious attention can be diverted from the auditory display, and will instinctively be returned to the display when significant changes or new events occur in the display data.

Parallel listening: It is possible to keep track of several streams of audio information simultaneously.

Acute temporal resolution: Audio perception involves high temporal resolution and range compared to other modalities.

Affective response: Audio perception has a strong ability to convey and influence affect (see Section 6.2).

Auditory gestalt formation: When multiple data sources are perceived as audio, overall patterns and trends can be perceived in the data.

Kramer also points out that combining audio display with other modalities can have extra benefits such as engagement, intermodal correlations, realism and synesthesia [54].

Disadvantages of Auditory Display for SBFS use [54:14]:

Low resolution of certain auditory variables: Parameters such as brightness convey only a limited number of differentiable values.

Lack of absolute values: Most auditory display schemes do not indicate the absolute values of the original data (one exception occurs when representing timing data with audio timing).

Lack of orthogonality: A change in one audio parameter may influence the perceived value of another parameter.

Annoying: Unwanted or aesthetically displeasing audio can be particularly annoying, due to the strong affective response evoked by audio.

Absence of persistence/no printout: Random access to different time-segments of audio data is generally only possible using memory, and audio data may be harder to save or share with others.

User limitations: Some users will have below-average perception capabilities for certain display variables, as in the extreme case of 'tone deafness' (ie. lack of pitch perception).

Lack of Common Description Language: (An additional disadvantage identified by the present author.) It is generally difficult for users to refer to and describe audio perceptions using words, except in the case of users with extensive/similar musical training.

6.2 Communicating Emotion Through Music

In the previous discussion, audio parameters were discussed as purely abstract values which can be modulated to convey information. However, music has strong connections with emotional state, and in the SBFS context, the information to be conveyed is information about emotional state, meant to be understood by the users (S→Cn) and/or to evoke emotional states in the users (S→Un). Therefore, the relationship between music and emotion must be closely examined in order to understand how best to manipulate audio parameters in the SBFS. A particular concern is that the strong pre-existing emotional associations of various musical features can lead to the accidental communication of unintended or erroneous information.

Affective/Metaphorical Audio Associations

Kramer describes two types of pre-existing association which can be triggered by various modulations in audio: *Metaphorical association* and *Affective association* [63]. Metaphorical association is "the association of a physical world variable with a metaphorically related change in an auditory variable."[63:212] For example, ascending pitch is often associated with upward motion, faster speed, or increase in a physical quantity. Affective association "concerns the linking of a user's subjective affect elicited by the emitted sound with related modifications in affect aroused by meaningful changes in the data." [63:212] For instance, if a sound becomes harsh or unpleasant, that may imply that a harsh or unpleasant event has triggered a change in the source data.

These two types of association cannot be identified immediately with separate pathways on the signal-flow model, since the unconscious and conscious may both process concepts such as 'more/less'. However, affective association seems more relevant to S→Un than to S→Cn. A key concern for SBFS design is that once a mapping is made from physiological input to audio output, changes in the input variables should evoke suitable affective and metaphoric associations. A notable example is that an audio signal which is jumpy or discontinuous in time often conveys associations of unease or annoyance. Therefore if jumpy, discontinuous data is not intended to convey an emotional state of unease/annoyance, the input-output mapping must be designed to smooth or minimize the discontinuity.

The two types of association described by Kramer involve changes or modulations in data, however there are also associations linked to the long-term or steady-state parameters of a given audio stream. For example, the author has observed that timbres with mainly low frequencies tend to be associated with the lower body and chest, and tend to be more soothing, while high-frequency timbres are associated with the upper body and tend to be more arousing.

Another type of steady-state association is triggered by the use of *representational* sounds. If a sound is similar to a real-world sound such as a dog barking or door opening, all the metaphorical associations of the real-world event will be triggered. Representational sounds can be employed in SBFS's to facilitate correct attribution of data, or to encourage specific affective associations. Representational sounds have not been explored extensively in the present work, other than the use of heartbeat-like sounds to represent heartbeat information.

Studies of Musical Communication Using Three Emotion-Theory Perspectives

A significant amount of research has attempted to explain the manner in which music communicates emotion. In linking musical variables to their emotional results, researchers have drawn on the three perspectives related to emotion theories discussed in Sections 4.1 and 5.1: the categorical perspective, the dimensional perspective, and the sentic perspective.

Music-Emotion Findings using the Categorical Perspective

Many studies have examined the musical techniques/variables employed by composers to convey emotion, by taking the categorical descriptions of emotion such as happy/sad/etc and matching these to a categorized musical 'features'. A useful summary of such findings is found in [14].

Juslin identifies examples showing that both composers and performers are able to intentionally convey classified emotional states to music listeners, with statistical significance [14]: Thompson and Robitaille asked musicians to compose short melodies indicating six basic emotions; when these pieces were played in a rigid, straightforward manner by a computer sequencer (ie. adding no performative emotional expression), listeners recognized the intended emotions [71]. A meta-analysis of performer emotional communication showed that performers successfully conveyed five emotion categories with an accuracy score of $p_c = 0.70$ [72].

As an example of more complex analysis still based on categories and features, Juslin and Lindstrom propose an Expanded Lens model of emotion communication in music, which allows multiple regression analysis to be employed in determining relative contributions and interactions between musical features [14].

Music-Emotion Findings Using the Dimensional Perspective

The literature does not suggest a 1:1 mapping between three particular musical dimensions and the emotional dimensions of valence/arousal/goal-state; however it may be that stronger correlations exist between certain pairings of musical and emotional dimension.

As Scherer and Zentner identify, studies have suggested a link between the musical dimension of consonance/dissonance and the emotional dimension of valence [73]. (For example Zentner and Kegan demonstrated a preference for consonant melodies over dissonant melodies in infants [74], while Borchgrevink demonstrated a preference for consonant chords over dissonant chords in albino rats [75].)

Evidence suggests that arousal is conveyed particularly effectively by tempo. For instance, re-interpreting from Juslin's summary chart of musical features [14:96], "fast tempo" is indicated for the categories of Happiness, Anger, and Fear, while "slow tempo" is indicated for Sadness and Tenderness; if placed on an arousal scale, the former three would rank higher than the latter two. (The chart also supports the valence/consonance link.)

The theories of Meyer look at music in terms of expectation and surprise conveyed by structure and variation [68]. This may support a link between goal-state and higher-level structural manipulation in music.

<u>Music-Emotion Findings Using the Sentic Perspective</u>

In contrast to the study of musical communication through category/feature based methods is the study of 'vitality affects'. Sloboda and Juslin outline how researches including Stern, Langer and Imberty have described music as containing dynamic contours which express "abstract 'forms' of feeling that occur both together with, and in the absence of, proper emotions." [24:79]. A connection can be seen between such theories and the 'essentic forms' of Clynes mentioned in Section 5.1. In fact, in one study by DeVries, subjects used a sentograph to translate the dynamic contours they felt in various pieces of music, and the pieces were found to have different characteristic contour patterns depending on their general emotional theme, similar to Clynes' sentograms for those emotions [76].

Emotion Representation versus Emotion Induction

Sloboda and Juslin point out that a dichotomy widely found in the music-emotion literature is that music "may both 'represent' emotions (that are perceived by the listener) and 'induce' emotions (that are felt by the listener)" [77:455]. They go on to state that this distinction is little-investigated and still unclear. It can be seen that this under-explored dichotomy relates to the question of relative importance of S→Cn versus S→Un in SBFS design.

Interestingly, this dichotomy is also directly reminiscent of the distinction between cognitive empathy and affective empathy discussed in Section 5.1. Perhaps music is sometimes internally modeled as an external system, having its own emotional state separate from the perceiver (pathway S→Cn), and sometimes experienced more directly, as the induced emotions are identified with the self (pathway S→Un) (see Figure 6-1).

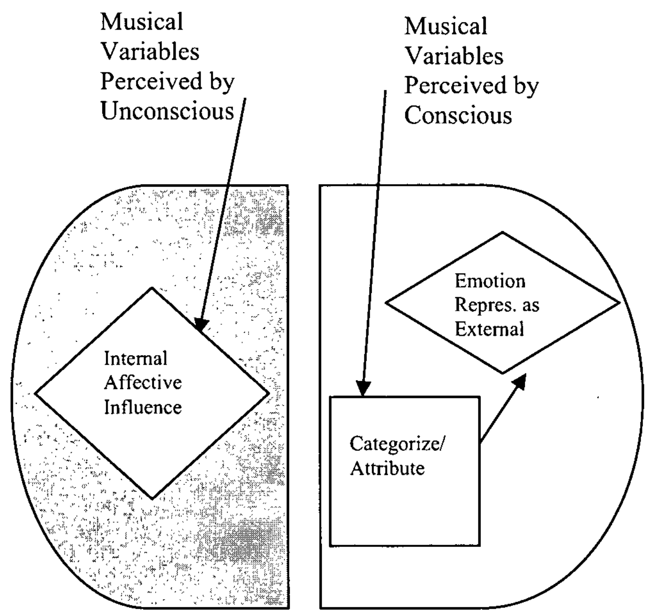


Figure 6-1. Identifying with Music vs. Modeling Music. Here the modeling/ identification distinction for music perception is shown using the same structure as Figure 5-3.

This possibility of identification with musical emotion raises the question for SBFS design, if User 1 experiences self-feedback (V1$\rightarrow$E1$\rightarrow$S$\rightarrow$U1), along with incoming emotional information from User 2 (V2$\rightarrow$E2$\rightarrow$S$\rightarrow$U1), will User 1 necessarily differentiate the two sets of emotion as 'self' and 'other'; or, will User 1 identify with both sets of emotions as being her own? This differentiation may be difficult unless aspects of the audio information are consciously perceived.

Also relevant to the topic of conscious/unconscious distinctions in music perception, the aforementioned studies of consonance and dissonance in infants and rats [74] [75] suggest to the author that musical consonance, and the emotional effects resulting from it, can be perceived along pathway S$\rightarrow$Un. Likewise, it is speculated that the more intricate or high-level musical structures (for example, the compositional technique of variations on a theme) depend more on conscious pathway S$\rightarrow$Cn to be appreciated.

Composition versus Performance

In many musical styles, the role of **composer** is separate from the role of **performer** (although the roles are often performed by the same person, but at different times, as in the case of the singer/songwriter). A SBFS may be designed to perform both roles simultaneously, in real-time. However, the composer/performer distinction yields useful knowledge regarding sets of parameter modulations which are able to convey emotion.

Typically a composer designs the musical parameters of pitch sequence, macro-scale rhythm, rough overall tempo, rough loudness level for different sections, rough timbre selection (instrument selection), and high-level harmony and structural progression. A performer manipulates note-by-note volume, small note-by-note timing offsets, rises and falls in tempo, note-by-note timbre modulation, envelope characteristics, and small pitch variations. The work by Juslin previously cited indicates that both composer and performers are able to accurately communicate emotion [14]; thus it can be concluded that both groups of audio manipulation parameters have the capability of communicating emotion.

6.3 Measuring the Impact of Music on Physiology

Some research on music and emotion has explored one step further along the signal-flow pathway S→Un→Vn by measuring the impact of music on physiology, which has direct relevance for SBFS design. Some of the notable conclusions from this research are listed here.

The pathway S→Un→Vn for music has measurable psychophysiological effects

In a meta-analysis of music psychophysiology, Bartlett found that in about 61% of studies undertaken from 1886-1996, ANS effects such as heart rate, muscle tension, and skin temperature were found to agree with the study hypotheses [78]. Khalfa et al found that GSR responses distinguish "sad or peaceful" music (ie. low arousal) from "happy or fearful" music (ie. high arousal) although not within these categories [79]. Harrer and Harrer showed that for a conductor, moments of greatest ANS response corresponded with passages judged to be emotional, not with passages of greatest physical exertion [80]. On the other hand, Katayose et al found that skin conductance patterns of musician and audience did not correlate well, nor did they correlate well with subjective scoring of musical affect [81].

The pathway S→Un→Vn for music can have measurable effects on heartbeat patterns

Harrer and Harrer found that syncopated rhythms and surprising musical events can cause premature heartbeats, that changes in volume and rhythm can cause parallel changes in HR, and that some subjects exhibit heartbeat entrainment with musical beats at times [80]. Studies have found that music can either increase or decrease heart rate depending on style and tempo [82] [83]; however some researchers have observed that music in general is more likely to increase than decrease heart rate [80] [84]. Because a consistent classification scheme for musical style has not been employed across the various studies, it is difficult to generalize results of these studies to make predictions about the heart rate effect of a novel musical style employed by a given SBFS. If music does have an overall tendancy to increase heart rate, it could be due to innate psychological factors or due to a prefererence of composers/performers for expressing arousal.

Conscious attentional stance can influence the strength of physiological responses, and physiological responses to music can occur unconsciously

Harrer and Harrer found that autonomic changes in response to music are greater when the subject is 'involved' than when the subject is 'critically analyzing', and that music may lead to autonomic reaction when the sounds are not consciously perceived (for example, background music or music played to sleeeping subjects) [80]. This suggests once again, as discussed in Chapter 4, that the pathway S→Un may be most effective when pathway S→Cn is weak.

Different users may have different psychophysiological responses to music

Harrer and Harrer found that relative response strength among different ANS measures (cardiovascular/respiratory/electrodermal) varies between different subjects and different musical pieces [80].

6.4 A Multi-Channel Representation for Display Pathways in the Signal-Flow Model

The proper construction of channels S→Un , S→Cn , SS→Un , and SS→Cn, as defined in the signal-flow model, is crucial to SBFS design. It is helpful to expand the level of detail of the model in this area so that various parameters used in system display can be identified. In the model, each single arc can be replaced with a bundle of parallel channels. As discussed further below, *each channel represents an independent parameter of audio generation/manipulation*, and *each channel can be described by a set of display-channel characteristics (DCCs).*

Figure 6-2 shows this configuration.

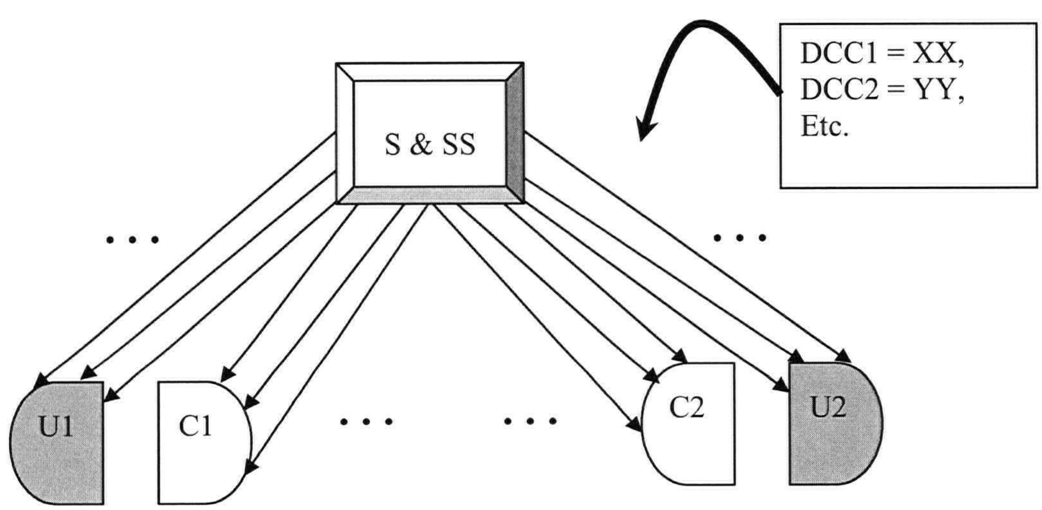


Figure 6-2. SBFS Display Channels. Each Channel *Represents* an Audio Modulation Parameter; Each Channel *Is Described By* perceptual Display-Channel-Characteristics (DCCs)

Display Channel Characteristics

The signals in each display channel begin as audio generation parameters which are well defined and chosen by the system designer, then transition through several different mediums (digital, analog and sound waveforms) before resulting in a complex set of perceptual variables in the user's brain.

The intermediate waveform representations are not directly of interest, so the mapping into perceptual space is treated by discussing *display-channel characteristics*: perceptual qualities which influence the transmission/perception of signal data, and have different expected values depending on the designed dimension of audio manipulation.

Based on the study of audio/music display, a basic list of DCCs has been identified as being relevant to the perception of display signals in a social biofeedback context. (It is not claimed that these characteristics are orthogonal or exhaustive, and it has not been determined how well they apply to non-audio display modalities.)

Basic Display Channel Characteristics

1. **dynamic range/precision**- the number of noticeably different values able to be taken on by the channel.
2. **perceptual accuracy**- the degree to which different values taken by a channel are able to be accurately ranked and compared.
3. **directional and absolute bias**- the additional (perhaps unintended) non-signal information which may be conveyed by a given pathway due to connotation. This bias results from both affective associations and metaphorical associations as described in Section 6.2 ; directional bias describes the connotations of *parameter change in a given direction,* while absolute bias describes the connotations created by *the perception of that varying parameter in general.* The value of each bias type depends on whether the target of the display channel is the conscious or unconscious mind. (Affective bias may tend to have a greater impact in channels

to the unconscious.) Bias can help or hinder the user in making or learning the correct links between audio generation parameters and emotional/physiological variables, depending on the goodness of fit between bias and emotional variable.

4. **intensity**- the degree to which a channel is noticeable and likely to draw mental perceptual/processing resources (potentially at the expense of other parallel channels). As discussed in Section 4.1, conscious attention is manipulated by both the conscious and unconscious mind-- sudden, affectively intense stimuli may gain the focus of conscious attention. Likewise, unconscious intensity can be damped once a channel gains conscious attention. Thus, for a given display parameter there is a complex relationship between intensity values for S/SS→Cn and for S/SS→Un.

It must be emphasized that for a given audio parameter, each of these characteristics may vary *depending on whether the destination of the display channel is the conscious mind Cn or unconscious mind Un.* While the audio-generation/manipulation parameters are independent, the DCCs are not independent- for example, conscious/unconscious associations may have an impact on intensity, and non-intense variables may draw fewer mental processing resources, thus losing precision and accuracy.

The display-channel representation thus attempts to separate **fixed characteristics** of a given display-mapping with the **dynamic information** flowing along that pathway. This separation is not absolute, however: an extremely important SBFS design consideration is that the behaviour of input signals (e.g. their range, mean value, variance and smoothness) will influence their perceptual precision, accuracy, bias and intensity. Thus, **when working out an approximate list of DCCs for a given combination of data input and audio display parameters, the expected dynamic behaviour of the input data must be taken into account.** As discussed below, signal behaviour can also be modified or re-mapped with signal processing techniques.

As a useful aid to SBFS design, a chart could be created from various research findings, linking possible input-type/audio-parameter mappings with the four DCCs, for the two cases of S→Cn and S→Un. Due to the complexity and extent of the literature as reviewed, such a chart has not yet been produced. Instead, the DCC concept can currently be used in system design by reviewing literature results relevant to every potential input/display mapping under consideration, and assigning approximate DCCs to each mapping.

Memory/Learning in Sub-Channels

The user's perception of SBFS display signals may be influenced by stored information about previous signals. In engineering terms, the perception of SBFS signals is not a memoryless process. The implications of this can be described with two additional display channel characteristics: memorability and learnability.

Additional Display Channel Characteristics Involving Memory

5. **memorability**- the degree to which a channel's character facilitates the recall at a later time of data conveyed by that channel.
6. **learnability**- the ease and speed with which signals in a given channel come to be fully/optimally perceived.

Memorability can have different values depending on the timescale involved, and it influences the ability to compare two signal values at different times, and therefore the ability to perceive patterns or correlations within or between signals. Learning, also, is too complex to be fully described by a single value, since it involves a cycle of learning and retention/forgetting.

Memorability and learnability are also strongly influenced by conscious/unconscious bias and the other DCCs.

Presenting memory/learning as attributes of the display channel is only one perspective on the underlying process: a more explicit view might include a processing element with memory as part of a transformation function located in each of Un and Cn. However, the DCC representation contains a useful level of abstraction for SBFS design.

Interaction Between Display Channels

One type of interaction between display channels that has been discussed is competition for attentional bandwidth, with reference to the DCC of intensity. Several other interaction effects can be described by *inter-display-channel characteristics* considered pairwise between display channels.

Inter-Display-Channel Characteristics

1. **grouping strength/stream identity**- the propensity of the channels to be perceived as belonging to a single 'stream' having a 'common source'.
2. **interference**- the amount by which the presence of a certain channel interferes with the accuracy of perceiving another channel.
3. **comparability**- the ease with which two channels can be mentally compared and patterns between them perceived.

Designing related parameters into the same stream can create the perception that they come from a 'common source'. Mapping two data sources into similar parameters in different streams can aid comparability, and parameters in separate streams may be less likely to interfere with each other.

Inter-channel interference is distinct from competition for attentional bandwidth. It refers to lower-level perceptual conflicts which stem from the mechanisms of processing particular audio variables. For example, drastic changes in the pitch of a tone may make it difficult to follow modulations in the brightness of that tone. Another example is that two pitch melodies in a similar frequency range may be harder to correctly disentangle.

Adjusting Bias by Shaping Free Audio Variables

Conscious and unconscious bias can be significant factors in SBFS design. One technique for ensuring an appropriate bias is to select appropriate audio generation parameters. However there is also the possibility of shaping bias through the manipulation of sound attributes not directly employed as information-carrying variables (which can be called 'free audio variables').

In particular, several display parameters will often relate to manipulations of a single perceptual stream whereas the timbre of this stream is not completely specified. Thus the timbre can be designed to invest the tone with extra bias attributes as desired. For example, if it is necessary to manipulate pitch, volume and tempo of a particular audio stream, the design choice is still available to give that stream a bright trumpet-like timbre, or the timbre of a muted female choir- each choice results in different affective and metaphor bias.

Adjusting DCCs through Signal Processing

It was shown in Chapters 4 and 5 that psychophysiological measures contain patterns at a variety of time scales and frequencies, and shown in this chapter that patterns in audio signals can be processed and perceived in markedly different ways depending on their time scale and behaviour. Therefore it is useful for system design to have techniques which can adjust the behaviour of the display signals to optimally convey the desired information.

Scaling/Remapping

It is important for both correct bias and accurate perception that the range of an input signal is mapped to an appropriate range of audio output. Thus input signals can be scaled and remapped by the SBFS (with linear or non-linear mapping curves) to ensure the desired output behaviour. For instance, a physiological change considered 'small' should not cause an audio change that will be perceived as 'large'. Another example is that input-to-pitch mappings will have better accuracy if output pitch stays in the central range of human perception.

Filtering

An obvious technique for adjusting frequency range/content is the use of analog or digital filters. These can be employed at several stages in the SBFS, from conditioning raw sensor data to remove noise, to smoothing audio parameter changes to avoid undesired startle responses. Conceptually, a single stream of physiological data such as IHR can be divided into several isolated psychophysiological parameters of interest by passing the stream through several parallel bandpass filters designed to pass different regions of signal behaviour. Dividing the signals in such as way can allow enhanced pattern perception compared to leaving the signals and frequencies intermingled.

Time Scaling/Looping

A technique for adjusting the time scale of a signal is to buffer a window of incoming data, then traverse the buffer for display at a more rapid rate than the original sampling acquisition rate. Once the end of the buffer is reached, data continues to be read from the beginning. This allows patterns of data to be presented at altered time scales. Such techniques will be referred to as 'sliding/looping window' techniques. Sliding/looping can be combined with nesting as discussed in Section 6.1. Note that this technique can produce a discontinuity at the loop point, which may be dealt with by smoothing or a pause/gap in output.

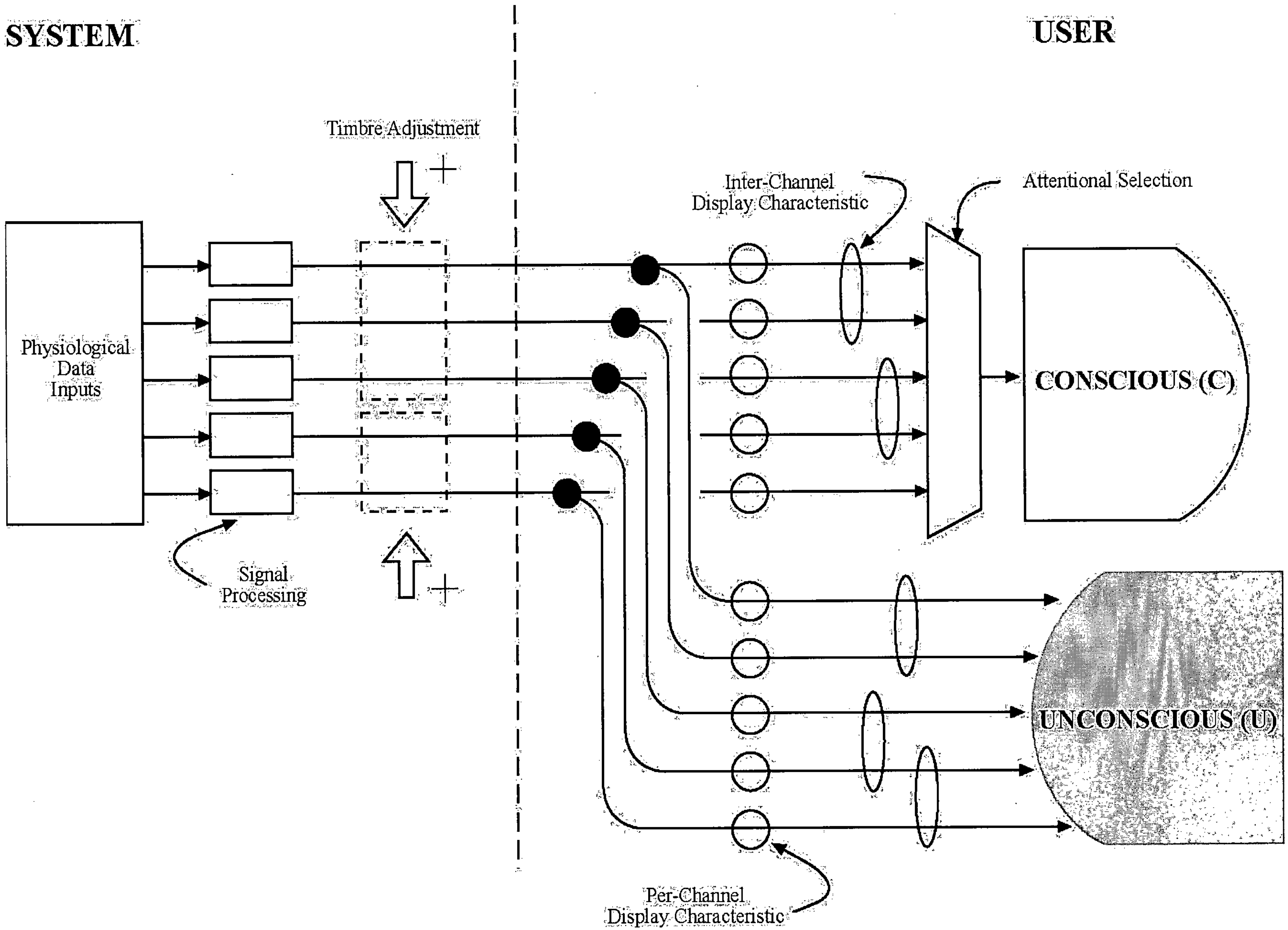
The looping of signal information is particularly relevant for the audio display modality since looping and repetition is frequently used in many styles of music. Depending on the level of time compression, sliding/looping can lead to perception of data in terms of different musical variables, from rhythm to texture to waveform/timbre. Looping can also allow perception/attention mechanisms to time-multiplex their focus back and forth to the loop/conversation.

An objection that could be raised against time-scaling is that a given contour of physiologically-based data might be optimally perceived/interpreted at its original, 'natural' time scale. However, Clynes' assertion that essentic forms can be perceived despite changes in modality or spatial scale suggests it is possible that time-scaling may not destroy the affective properties of the signal [45].

Summary Diagram: The Detailed Display-Channel Representation

Figure 6-3 (next page) shows how the concepts of display channels, DCCs, stream grouping, signal processing and bias adjustment can be shown within the framework of the signal-flow model.

Figure 6-3. Detailed Design and Characteristics of SBFS Display Channels.


SYSTEM
USER
Timbre Adjustment
Physiological
Data
Inputs
Signal
Processing
Inter-Channel
Display Characteristic
Attentional Selection
CONSCIOUS (C)
UNCONSCIOUS (U)
Per-Channel
Display Characteristic

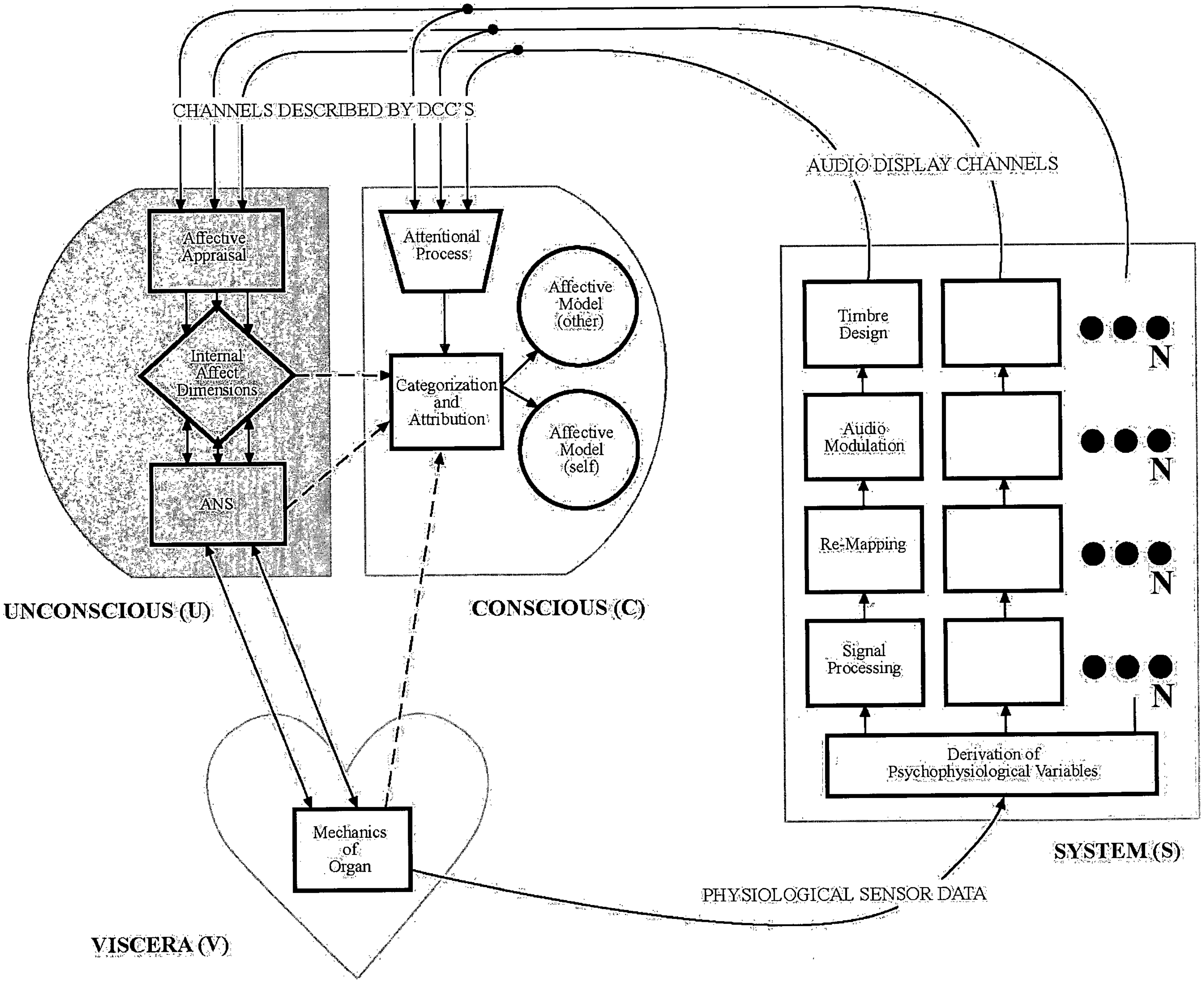
7. Summary of Common Themes in the Signal-Flow Model and Principles for SBFS Design

Reviewing the literature relevant to different regions of the signal-flow model has enabled many details to be added to the model and several relevant phenomena and themes to be identified. This section shows how the collected information can be considered across the various regions to gather useful themes and design concepts which apply to the model as a whole.

7.1 A Detailed Diagram of SBFS Input and Output

Figure 7-1 (next page) integrates many of the details added to the signal-flow model in Chapters 4-6. It highlights how detailed within-user affective channels relate to the detailed channels of the artificial system. The diagram focuses on the information most useful to the system designer, by only showing connections between one user and the system-- system connections to the other user are identical in nature and can be imagined as a mirror image. User-user natural communication channels (see Chapters 3 and 5) are not pictured but are to be kept in mind. Considerations specific to the SS node are not included in this diagram, but are summarized in the following section.

Figure 7-1. Detailed Diagram of SBFS Input and Output.



CHANNELS DESCRIBED BY DCC'S
AUDIO DISPLAY CHANNELS
Affective Appraisal
Internal Affect Dimensions
ANS
Attentional Process
Categorization and Attribution
Affective Model (other)
Affective Model (self)
Timbre Design
Audio Modulation
Re-Mapping
Signal Processing
N
N
N
N
Derivation of Psychophysiological Variables
UNCONSCIOUS (U)
CONSCIOUS (C)
SYSTEM (S)
Mechanics of Organ
PHYSIOLOGICAL SENSOR DATA
VISCERA (V)

7.2 A High-Level Strategy for SBFS Design

Developing the signal-flow model has illuminated the concerns and principles of SBFS design. By summarizing the findings of Chapters 3-6, and by considering the complete model in its entirety, the following integrated perspective has been created:

General Design Principles

1. The fundamental design goal of a SBFS can be described as conveying affect information from the unconscious or conscious mind of one or more users, to the conscious or unconscious mind of one or more users, according to 'SBFS paradigms' such as those listed in Section 3.3.

2. Basic inspirations or guiding examples of how to transmit affective information can be found by considering natural within-person affect transmission/feedback processes and natural between-person affect communication processes. Either of these two process types can be employed as guidance; in fact there is similarity between the two process types.

3. Based on the study of natural affect transmission processes, it is proposed that the type of affect information to be transmitted from and to the unconscious mind can be described in terms of 'dimensional affect variables', and further proposed that the type of affect information transmitted from and to the conscious mind can be described in terms of 'categorical affect representations'. It is also speculated that various other levels of affect representation within the mind (such as autonomic nervous system variables or the physical variables of viscera) can be interpreted by the brain into affect information.

4. The task of artificially transmitting affect information can be logically divided into two parts: designing data input/derivation from biosensor data, and designing a mapping from input into output/display.

System Input Design Principles

5. A promising method for deriving the 'categorical affect representation' from biosensor data is to employ machine-learning/statistical techniques, or analyses based on previous research using such techniques.

6. One method for deriving the 'dimensional affect variables' from biosensor data is to use measures and findings from the field of psychophysiology; however these findings do not yet give an exhaustive portrayal of the 'dimensional affect variables'.

7. Another method for deriving the 'dimensional affect variables' is to discover biosensor/dimensional-var. relationships through experiment or trial and error—SBFS's can be useful exploratory tools in this regard.

System Output Design Principles

8. Various findings have shown that audio/visual modalities are capable of conveying affect information to the conscious and unconscious mind.

9. Not every audio/visual display parameter is equally effective at conveying given affect information. Display parameters vary in terms of perceptual acuity, suitability for a given affect dimension, and tendency toward conscious/unconscious perception. The varying qualities of display parameters can be usefully examined in terms of the 'display-channel characteristics' (DCCs) suggested in Chapter 6.

10. The DCCs of a desired input/display mapping can be adjusted through timbre design, and through signal-processing methods such as filtering, looping, and re-mapping.

11. One method for determining the DCCs of a particular input/display mapping is to use findings from auditory display research; however these findings do not yet give an exhaustive set of mappings vs. DCCs.

12. Another method for determining the DCCs of a particular input/display mapping is to study musical composition and the works of composers. (As a side note, **music itself** could in a sense be characterized as a social biofeedback system!)

13. Another method for determining the DCCs of a particular input/display mapping is through experiment/trial-and-error.

14. Given the qualitative/perceptual nature of DCCs, their dependence on source data behaviour, and the absence of an easily-accessible/comprehensive literature in this regard, SBFS design should employ a large degree of user testing, and in-design testing.

Miscellaneous Design Principles

15. Because SBFS's co-exist with and augment natural, rich communication pathways, interference between signals is a design concern. Three types of potential interference are: perceptual interference between display parameters, interference between display parameters and/or social communication due to attention competition, and interference between system physical aspects and the social environment.

16. Identification and modeling are important concepts for SBFS design: an open question regarding SBFS design is whether/how to encourage the correct attribution of incoming affect information.

17. Learning and memory effects mean that transmission path strengths in a given SBFS with given users can vary over time.

18. Multi-user joint physiological measures may be useful to employ in SBFS design; however they are not well-understood, and natural guiding-examples for transmitting this information have not yet been identified. Correlation measurements between two users' physiological data streams may yield a measure of empathy.

7.3 Suggested Social Biofeedback System Design Cycle

The following list of design steps has been developed as a first attempt to formalize the SBFS design procedure. It is suggested that performing the design steps in approximately this order can help prevent design decisions from conflicting with each other.

Order within each phase should only be loosely followed- interaction between different SBFS components necessitates considering many factors in parallel. For example, sensor design chosen influences input variables possible and vice versa.

These steps were developed partly from experience gained in the design and testing of the 2Hearts system (see Chapters 9-12) and partly from the theoretical findings relating to the signal-flow model.

<u>PHASE I – System Goals</u>

1. Describe high-level system goals/intent and frame goals in terms of single or combined SBFS paradigms (see Chapter 3)
2. Define explicit goals and performance measures based on the relevant SBFS paradigm(s)

<u>PHASE II – System Inputs</u>

3. Select physiological measures Vn→S and Vn→SS based on paradigm and psychophysiological measures theory (see Chapters 3 and 4)

4. Select sensor type capable of measuring desired physiological data (see Chapter 8)
5. Define system physical design based on sensor constraints, keeping in mind interactions with social-environment (see Chapter 5)
6. Design physiological data acquisition/pre-processing scheme (see Chapter 4)

PHASE III – System Outputs (see Chapter 6 regarding each step)

For each physiological variable:

7. Design filtering and looping algorithms if required
8. Group variables into a perceptual stream with respect to inter-display-channel characteristics
9. Select target audio modulation parameter(s) considering associations, interference and other display-channel characteristics
10. Design re-mapping and range warping algorithms to maximize precision/accuracy/correct bias, using recorded sample physiological signals as reference

For each perceptual stream group:

11. Employ unmapped audio variables to perform timbre design and bias adjustment, using recorded sample physiological signals as reference inputs

PHASE IV – Testing

12. Test each audio-modulation display parameter separately, and in combination, by measuring effects on user physiology and user experience
13. Iterate by returning to phase II if necessary, until good parameter design is validated
14. Test complete system in social context, by measuring effects on user physiology and user experience with respect to paradigm goals
15. Iterate by returning to phase I, II or III as appropriate

The iteration listed in step 13 has been found to be especially important. Experience suggests that the 2Hearts design cycle could have been quickened and improved by earlier/more testing of this type.

Note that this design cycle presents a high-level description of the design/testing process. The details of each step must be elaborated depending on the particular system design. As various SBFS paradigms are explored further in future work, it will be possible to list suggested performance measures and testing methods for each SBFS intent/goal paradigm. Chapters 11 and 12 include an initial exploration of one method for implementing Step 12 above, and also include testing methods relating to Step 14 for the case of the *control, converge* and *understand* paradigms.

8. Overview of Technologies Considered for 2Hearts System Design

Chapters 3 through 7 have described one major contribution of the current project: the theoretical SBFS signal-flow model. Here begins the discussion of the second major research contribution: development of the 2Hearts SBFS, intended to exemplify and further develop various aspects of the theoretical model.

As necessary background for the explanation of 2Hearts system design, this chapter describes the hardware and software component technologies that have been employed or investigated for the 2Hearts system.

8.1 Heart Sensor Technology

Several types of heart sensor have been investigated for use in the 2Hearts system:

Finger/Ear-Clip Bloodflow Sensor

This type of sensor detects varying blood volume in a finger or earlobe by passing infrared light through the body part and measuring how much light is absorbed. Output is an analog voltage (~1V p-p) which rises and falls for every heartbeat, according to blood flow behaviour. The model investigated was Heart Sensor HRM-DIN from Vernier [85], now discontinued. Figure 8-1 shows this sensor.

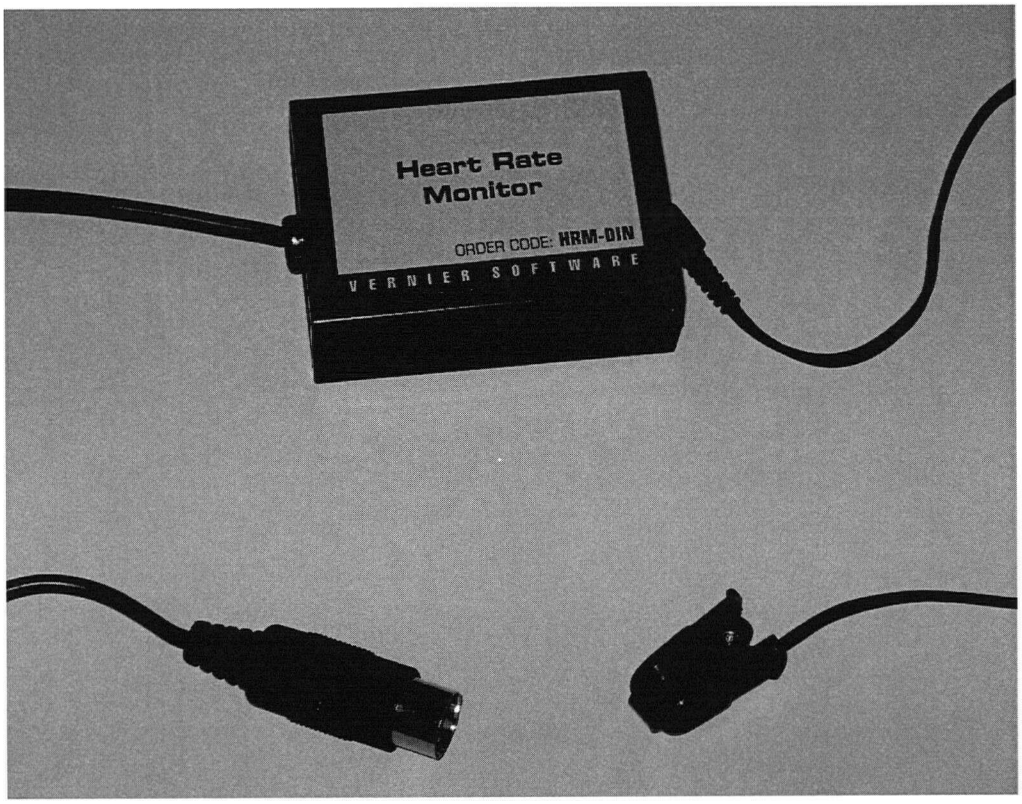


Figure 8-1. Finger-Clip Heart Sensor.

Advantages of this sensor is that it is easy and non-invasive for users to attach/detach, and it provides information about blood-pulse shape rather than just trigger data. Disadvantages are that the sensor must be manually calibrated for light conditions, that user motion causes surges of blood-flow and erroneous readings, and that the detected pulses are not sharp enough for accurate peak-detection.

Procomp+ ECG Sensor

This sensor package is a legacy model from Thought Technologies [86]. It includes various optional biosensors (such as temperature, skin conductance and ECG) and a digitizer module which reads up to 8 sensors in parallel and transmits sensor data to a PC serial port via an optically isolated link. ECG data is digitized with a maximum sampling rate of 256Hz and a precision of 13 bits.

Advantages of this sensor are that it supplies the entire PQRST heartbeat wave (allowing the use of psychophysiological measures based on the wave shape), and that it allows two chest electrodes and a separate ground electrode to be placed as desired, potentially reducing sensor noise compared to a two-electrode/fixed-position device such as the Polar sensor discussed below). A disadvantages is that the electrodes must be stuck to the skin of the chest with adhesive (which can be uncomfortable). At the time the technology was reviewed, a closed-source DLL was required to read the sensor data into custom software; this introduced buffering and latency issues.

Polar Chest-Strap Heart Sensor

This type of sensor, developed by the Polar company [87], works similarly to an ECG sensor by measuring the voltage difference between two points on the skin of the chest. Analog circuitry in the chest-strap performs peak detection on the electrical heart pulses, and transmits a short-range (~1m) wireless pulse signal to a receiver module, once per detected peak. The receiver outputs an analog voltage pulse (~3V p-p) once per received pulse signal. Output pulses are constant in shape. The model investigated was Exercise Heart Monitor EHR-DIN from Vernier [85]. Figure 8-2 shows this sensor.

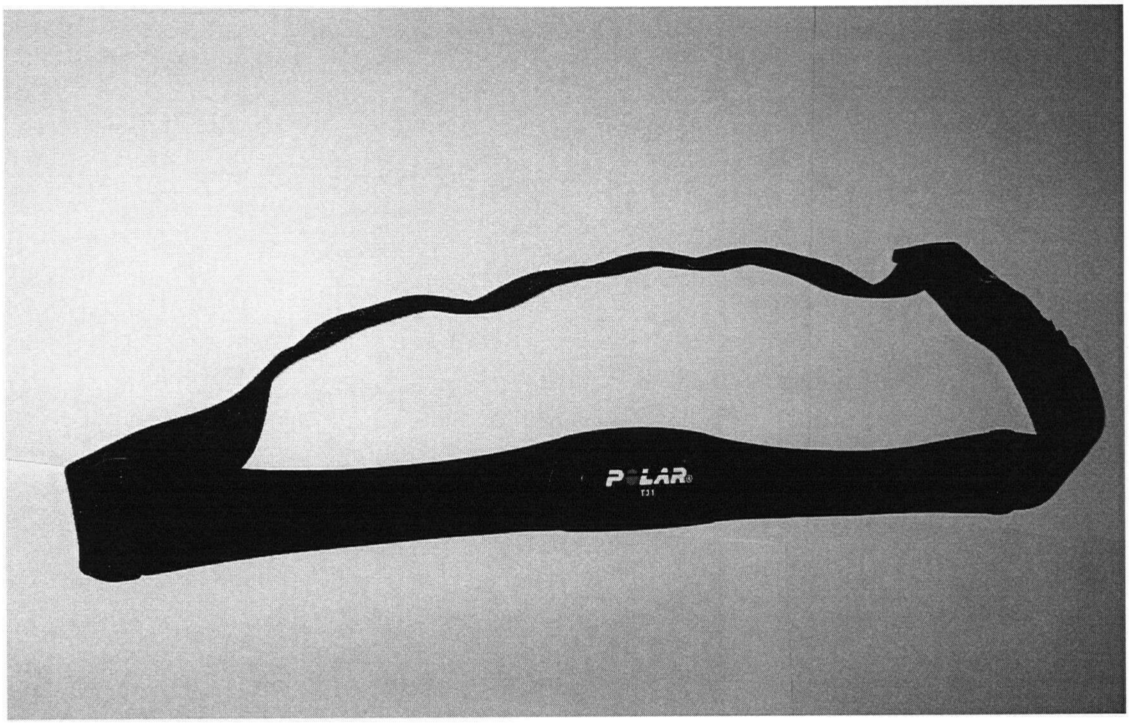


Figure 8-2. Polar Chest-Strap Heart Sensor. (The small receiver module is not shown.)

Advantages of this sensor type are the suitable connotations of detecting heartbeats from a chest-placed sensor, the good timing precision of detected pulses, and the resistance of these sensors to motion artifacts once the belt is snugly in place. Disadvantages are the short range of wireless transmission, the potential crosstalk between nearby transmitter/receiver pairs, and the somewhat intrusive user requirement of placing the sensor next to the chest skin. Section 10.4 describes how the sensors were modified to avoid crosstalk The suitability of the Polar sensor for SBFS use is supported by the prevalence of this sensor type in recent wearable computing research [88] [89].

BioMuse

This hardware package is intended for use in biosensor-based art and music. Developed in 1992, it supports up to 8 biosensor inputs which are digitized at 30kHz/12 bits, and outputs data as control messages in the MIDI protocol [90]. The BioMuse was not explored in detail for 2Hearts system use, due to its high cost (approx. $20, 000 US).

Future Sensor Types

An interesting future development for SBFS's is that the Vancouver—based company Wireless 2000 is currently developing non-invasive, remote wireless heartbeat sensing technology based on ultra-wide band electromagnetic pulse reflection [91]. This could allow SBFS in the future to perform heart sensing which would be very transparent to the social context.

8.2 Heart Signal Digitization

Given the analog signal output of the finger-clip and Polar sensor types, additional hardware is required to digitize the signals for computer processing. Three methods of digitizing heart sensor data have been employed in 2Hearts system designs: a custom built analogue-to-serial interface built on a prototype board using the PIC16C73 chip, the I-CubeX voltage-to-MIDI converter, and the Echo Indigo I/O PCMCIA audio interface.

The serial board required the use of custom-coded drivers to interface with audio software, while the I-CubeX outputs MIDI signals processed natively by most audio software. However, the implementation of MIDI in PCs creates the potential for timing jitter. (An informal loopback test, with data sent from PC to I-CubeX at regular intervals and re-digitized, indicated a combined output/input jitter of up to approximately 20ms). The Echo card has been found to be the best solution (see Section 10.4).

8.3 Data/Audio Processing Technology

Two freely-available audio-software frameworks have been used to build 2Hearts software: jMax [92] and Pure Data (PD) [93]. Each framework employs an object-and-connection based GUI to facilitate the creation of various audio/data processing algorithms. Each framework allows the creation of custom-coded processing objects using a C-language API. Each runs on Linux and on Windows. It was found that Pure Data crashed less often then jMax, and Pure Data runs from a convenient "one-folder" installation.

9. Findings from 2Hearts Pilot Systems

Initial exploration of 2Hearts design involved the creation of two concept systems, 2Hearts--Version 1 and 2Hearts--Version 2 [94]. These systems were not designed with a structured theoretical approach, but experiences with their design raised many of the issues later researched for the SBFS signal-flow model, including system goals/intentions, conscious/unconscious perception and attention, and associative bias of musical mapping strategies.

The central theme of the evolution from system to system was the exploration of different approaches to the musical display mapping. Version 1 used pre-composed music segments which transitioned based on momentary and cumulative heart-data measures. Version 2 explored a more direct approach by mapping heart rates to scaled pitches. Version 3 was then designed to combine valuable aspects of both approaches by layering several differently-audified heart-data-derived variables, in an attempt to encourage emergent rhythmic patterns to be perceived musically.

The following sections briefly summarize the design and findings of 2Hearts--Versions 1 and 2, followed by a discussion of graphical display techniques which were also prototyped.

9.1 Design and Findings of 2Hearts--Version 1

2Hearts—Version 1 employed finger-clip heart sensors, a serial port A/D interface, and software built in the jMax framework running under Linux on a Pentium II computer. The software controlled musical output by sending MIDI control signals to an external Yamaha EX5 synthesizer. Participants sat in two chairs with one hand each placed on the table holding the sensors.

In the Version 1 display mapping, each user's HR was represented by a different instrument playing a pre-composed looping score segment as part of a synchronized duet. Higher/lower per-user HR mapped to higher/lower pitch height through the selection of alternate pre-composed high/low chord inversions. Mean user-user heart rate controlled the overall tempo of the music, and the difference in heart rates (smoothed to give a slowly varying trend) controlled the volume of a third audio stream composed to portray 'conflict'.

High-level progress through different pre-composed score segments was according to a 'rolling ball' control metaphor [94]- users' heart rates were treated as virtual 'forces' along perpendicular axes which guided the progress of musical state between regions of a virtual terrain. Reasons for using such a control metaphor were: to allow variation in the music, to have a display variable indicating the long-term cumulative history of heart rates, and to employ inertial damping of control thus improving the long-term controllability of a musical variable.

The system was evalulated by several informal session where pairs of volunteers played with the system.

Notable findings from this system were:

- The motion-sensitivity of the fingerclip sensors detracted from the user experience by forcing users to either remain relatively still (and thus socially hampered) or cause erroneous heart data.
- Subjects reported enjoying the music and seemed excited or amused when hearing obvious changes in the music.
- Subjects did not seem to understand the HR-difference mapping.
- In general the music mappings were quantized too coarsely to perceive short-term HR patterns.
- The rolling-ball metaphor did not appear to be understood by users beyond the basic concept that the music would sometimes change. Probably the metaphor

added a level of indirectness and complexity to an already novel and indirectly-controlled system.

9.2 Design and Findings of 2Hearts--Version 2

2Hearts--Version 2 employed Polar heart sensors (modified as described in Section 10.4), an I-CubeX voltage-to-MIDI converter, and software built in the jMax framework running under Linux on a Pentium 4 computer. Participants stood inside a CAVE virtual projection environment and wore head trackers and shutter glasses; this allowed incorporation of graphical display as described in Section 9.3

The audio display of Version 2 primarily employed a simple mapping where each detected user heartbeat triggered a note, with note pitch controlled by HR and quantized to semitone values. The aim of this mapping was to explore the perception of oscillatory and tonic HR patterns when framed as melodic pitches.

Notable findings from this system were:

- The Polar sensors worked accurately and did not produce many motion artifacts
- The required VR wearable apparatus was cumbersome and prevented a feeling of natural social interaction between users
- The use of pitch in semitones sometimes successfully created random 'musical' sounding melodies
- No particular HR phenomena were identified as perceivable in terms of particular melodic characteristics
- The simple mapping grew harsh and annoying after a short period of use; in part due to a sound design for the notes which was too bright and percussive
- Subjects seemed able to intentionally achieve mutual states of high or low HR by exaggeratedly play-acting different social scenarios such as angry arguments

9.3 Discussion of Graphical Display Methods in Pilot System Designs

Two different schemes for graphical display of heart sensor data were prototyped in the course of 2Hearts system development; an 'aura' scheme which was part of 2Hearts--Version 2, and a 'waves' scheme which has not yet been incorporated into a particular finalized 2Hearts version.

This paper does not address in detail graphical display methods for SBFS use. However, the two graphical schemes attempted are outlined and some brief observations are made, because they compare and contrast with the audio display findings in interesting ways. (See also the discussion of auditory vs. other display modalities in Section 6.1).

'Aura' graphical display

The graphical display of 2Hearts--Version 2 used a CAVE projection environment and a split-screen graphics programming technique to create a separate 3D visualization from each user's perspective. Each user saw a cloud of 'aura' particles around the other, which emminated from the user position with each heartbeat and flowed outwards. Aura color was modulated by HR, with blue resulting from low values and red from high values. Difference or similarity between users' HR values was portrayed by an attracting/repelling behaviour of the two auras. The virtual 'ground' also changed height according to users' mean HR, in order to give a feeling of 'levitation' caused by joint arousal. Figure 9-1 shows the 'aura' display.

Figure 9-1. The Aura Graphical Display Scheme for 2Hearts.

The 'aura' display software was coded using the VRJuggler framework [95], executed on an SGI Onyx computer , and controlled by signals sent from the 2Hearts--Version 2 jMax software via a network connection.

'Waves' graphical display

This graphical display scheme uses an additive color-blending technique to visualize 3 horizontally-spreading wave-patterns corresponding to user 1 heart data, user 2 heart data, and combined heart data. The design employs the animation technique of repeatedly moving/copying pixel data from the central area toward the sides. (Apparent motion is slowed near the edges of the display to create a 'fisheye' effect.) For each detected heart trigger, a standard time-varying color-intensity pulse is painted onto the central buffer. User 1's area modulates from black to green, and user 2's area modulates from black to red. The centre area combines the red (R) and green (G) pixel-color information into a single color, thus creating shades of red, green, and yellow. In this way, the display is intended to portray phenomena within the per-user and inter-user data as distinctive color-band patterns. Figure 9-2 shows a screen image of this display scheme.

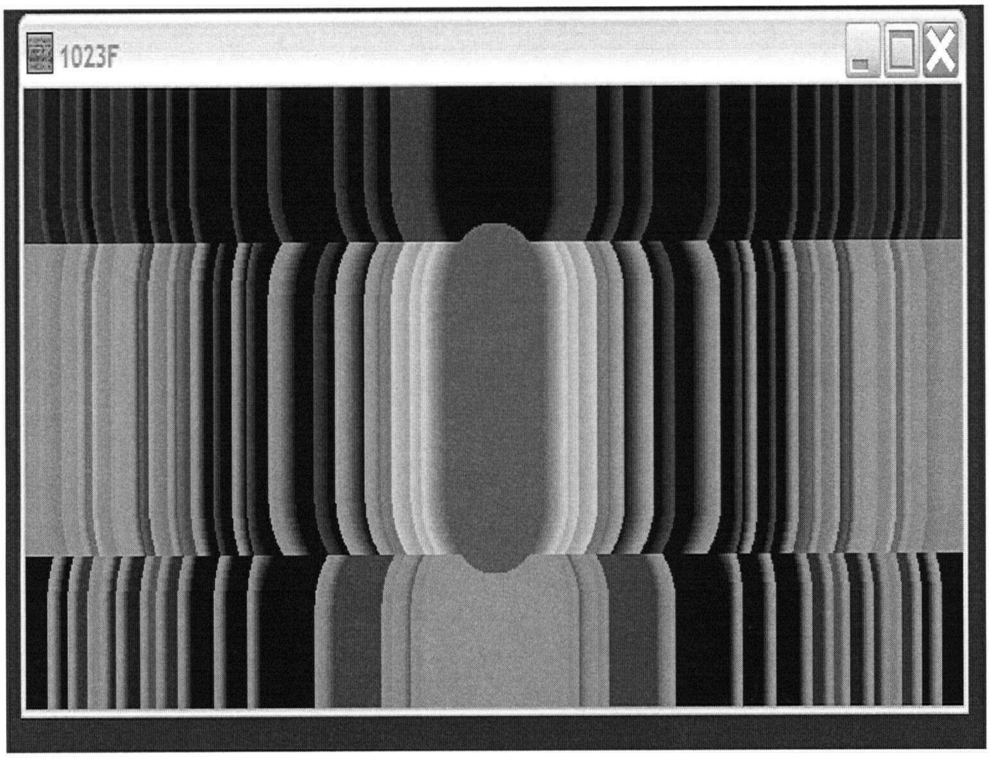


Figure 9-2. The Waves Graphical Display Scheme for 2Hearts. The upper portion ranges in color from black to red; the lower portion from black to green. The middle portion contains a mix of black, red, green and yellow.

This display is intended to be projected on a table sat at by the users or a wall parallel to the users. It was implemented within Pure Data using the 'Framestein' graphics extension.

Some interesting observations regarding the prototype display schemes are:

- During informal testing, users of the 'aura' system explicitly talked about the aura colors with each other; while they did not discuss the audio display. This is related to the 'Lack of Common Description Language' problem discussed in Section 6.1
- Because graphics are precisely located in space, it was easy to avoid direct perceptual conflict between system graphics and users' view of each other- auras were placed immediately around/beside users' faces, but not obscuring them. This is a benefit compared with audio display, as noted in the 'Lack of spatial information' point in Section 6.1.
- Again due to the spatial specificity of graphics, a strong identity connection was made between each user and their aura.
- The 'waves' scheme shows how a long history of source data can be perceived in parallel to look for trends. Looping techniques for audio achieve a similar goal, but not with the degree of true parallelism. This relates to the 'Absence of persistence' point in Section 6.1.
- In general, the explored graphical schemes seemed to highlight different types of patterns present in source data as compared to the explored audio schemes.

10. Design and Implementation of 2Hearts--Version 3

This section presents the design and implementation of the most recent version of the 2Hearts social biofeedback system. Most of the principles of the current SBFS signal-flow model and design cycle had been identified by the time that design took place. However, parts of the theory were refined as a result of experiences developing and working with this 2Hearts version. The presentation here follows a similar order to SBFS design cycle Phases I-III.

10.1 Design Phase I: System Purpose/Paradigm

2Hearts was conceived as a "two-person instrument", a system which encourages users to listen to the musical output and take active control in shaping this output by collaborating and finding social interaction strategies. This complex system goal fits the *co-navigate* paradigm described in Section 3.3.

The design implications shown by the co-navigate paradigm include many signal-flow model pathways to be augmented, and no pathways to be damped. This was considered appropriate for an exploratory system like 2Hearts, allowing many different pathways to be tested without the more challenging task of damping certain pathways while augmenting others. Also, the co-navigate paradigm contains elements of some of the simpler paradigms within it, including the *control* paradigm, the *converge* paradigm and the *understand* paradigm (minus the optional damping of the control paradigm). The user testing was designed to take advantage of this and analyze subsets of the system by testing the control and understand paradigms.

10.2 Design Phase II: System Inputs

Sensor Design

It was not within the scope of the project to design complex custom sensor hardware, so a heartbeat sensor was chosen from the available off-the-shelf technologies: infrared blood-flow sensors, exercise chest-strap sensors, or ECG sensors (see Section 8.1). It had already been determined that the clip-based bloodflow sensors caused significant problems with missed readings and motion artifacts (see Section 9.1), and the Polar sensors were more resistant to motion, with fewer erroneous readings. Polar sensors were chosen over ECG since they are less expensive, already shown to give acceptable results, and have a quicker/less intrusive attachment procedure compared to multiple stick-on electrodes. Polar sensors also have the benefit of encouraging attention to the chest area and thus to the heart. As noted in Section 10.4, the Polar sensors were used as 'wired' sensors due to a modification necessary to avoid crosstalk.

Physical/Environmental Context Design

Considering the arcs of the *co-navigate* paradigm with respect to the key SBFS concern of *interference* shows that the physical aspects of the system must be designed to interfere as little as possible with normal verbal/non-verbal communication (arcs Vn/En→Un/Cn). Combining this requirement with the constraint of wired sensors, the system was designed for users to be seated on a comfortable couch, with sensor cables running to a nearby laptop computer connected in turn to a home stereo system and a pair of speakers. The speakers are to be located in a natural listening position not too far from or close to the users. In this way, the natural movements of social interaction on a couch are not greatly impeded, the setting seems natural, and physical contact and line of sight between users is facilitated, but not forced. The use of wired sensors has been found to be relatively comfortable and acceptable, but developments in remote sensor technology may soon enable further-developed systems to interfere much less with natural body motion (see Section 8.1).

Psychophysiological Measures

As mentioned above, the Polar heart sensor technology constrains psychophysiology measurements to those derived from time-of-beat readings (rather than QRS pulse shapes or blood flow/pressure data). This was judged acceptable for the system since several interesting IHR-derived measures to explore were found in the psychophysiology literature (see Section 4.2).

It is noted in Section 7.2 that three major approaches to deriving psychophysiological measures are: using machine-learning techniques, drawing on identified psychophysiological phenomena, and exploring potential measures within the design process. For 2Hearts, the latter two approaches were employed, but not the first. One reason for not focusing on machine-learning techniques for emotion recognition is that these techniques are already being actively explored in the context of Affective Computing (while the other techniques are less-explored and may yield exciting new directions). Another attraction of the approaches employed is that they are suggested to relate more closely to the internal unconscious representation of affect; and transmission to the unconscious is one of the most unique characteristics of SBFS's (and expected to be more novel and challenging then transmission to the conscious mind).

In general, 2Hearts design was performed with particular concentration on the transmission of affect information to/from nodes Un, as shown in Figure 10-1 below. The reason for this focus is that transmission of unconscious information is an especially unique aspect of the theoretical work on SBFS's presented here, and hence deserving of detailed exploration.

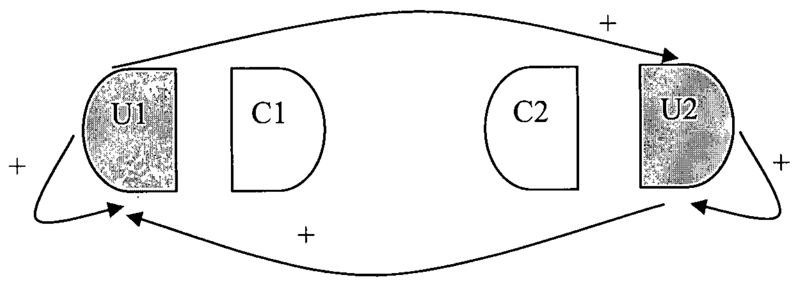


Figure 10-1. Design Focus of the 2Hearts System. Within the SBFS co-navigate paradigm, the arcs of aumentation shown here were judged to be especially interesting to explore in the context of a heartbeat/audio –based SBFS.

The *co-navigate* paradigm does not specify a focus on a particular set of within-user affect-related variables. The study of heartbeat psychophysiology in Section 4.2 found that in heart rate-based measures tend to be associated with the autonomic nervous system, and with the dimension of arousal. This fact placed a certain theme or constraint on how the *co-navigate* paradigm could be implemented—a heartbeat-sensor-based system may tend to highlight affect information relating to arousal.

It was found that IHR-derived measures fall under the categories of tonic HR, phasic HR medium-term elevation, phasic HR acceleration/deacceleration patterns, HRV measures related to the HR frequency power spectrum , and newer/experimental measures based on chaos, fractal, an entropy concepts.

For 2Hearts design, a general decision was made to focus on the simpler measures of IHR. As discussed, there is a lack of general consensus on how best to define and process HRV. Moreover, the various phenomena observed directly from the basic IHR signal are linked to many of the desired psychological categories of interest, including conscious-linked phenomena (orienting response and cognitive load), versus unconscious-linked phenomena (stress, anxiety, excitement, and physical pain); and goal-state (anticipation, appetitive motivation) versus arousal (most of the other measures). (Notably, the emotional dimension of valence has fewer well-defined connections with heart measurements in the literature.) This variety of phenomena directly related to HR means that a IHR-based set of psychophysiological measures can explore a cross-section of U/C→V→S pathways as desired (with the general emphasis on arousal as noted above).

In fact, all other HRV measures are derived from the IBI information, and such methods strive mainly to detect repeating patterns or chaotic behaviour in the heart-rate information. It is this type of analysis that is interesting to pass over to the human perceptual systems using an appropriate data- display scheme—such experimentation may lead to new insights into the patterns present in HR and HRV signals.

Specific IHR variables Chosen

As outlined in Section 7.2, deriving specific HR measures is an exercise in working 'back up the chain' to estimate various signals that might be present between the heart mechanics and the underlying/internal affect dimensions. The specific HR measures chosen to make available as inputs for the display mapping were time-of-beat for users 1 and 2, and IHR 1 and 2, derived from the interbeat intervals. Time-of-beat signals, clearly, convey information directly from the mechanical action of the heart which may be psychologically recognizable and relevant. On the other hand, the variable of IHR may act as a rough approximation of the ANS parameter of sympathetic action (again, a variable potentially recognizable as meaningful affect information). It was reasoned that other ANS phenomena observed as distinctive patterns in HR could be conveyed to the users as emergent patterns perceivable in the modulation of these basic signals. Whether or not these emergent patterns can be easily/naturally interpreted in a psychologically relevant way, was an open question to be explored.

This approach allowed the exploration of the most intriguing aspect of the heart-input/music-output combination: direct translation of psychologically relevant **rhythmic** patterns in the heartbeat to perceptible psychologically relevant **rhythmic** pattern in audio. This approach also avoided the need for complex or error-prone detection/

parameterization algorithms for specific HR patterns outlined in the psychophysiology literature.

Figure 10-2 shows how the IHR-derived measures employed in 2Hearts system design relate to within-user affect variables. Notice that the filtering stage of signal processing (described in Section 10.3) is also shown- due to the design goal of transmitting affect information both to and from users, the input and output stages of an SBFS such as 2Hearts are best considered in conjunction.

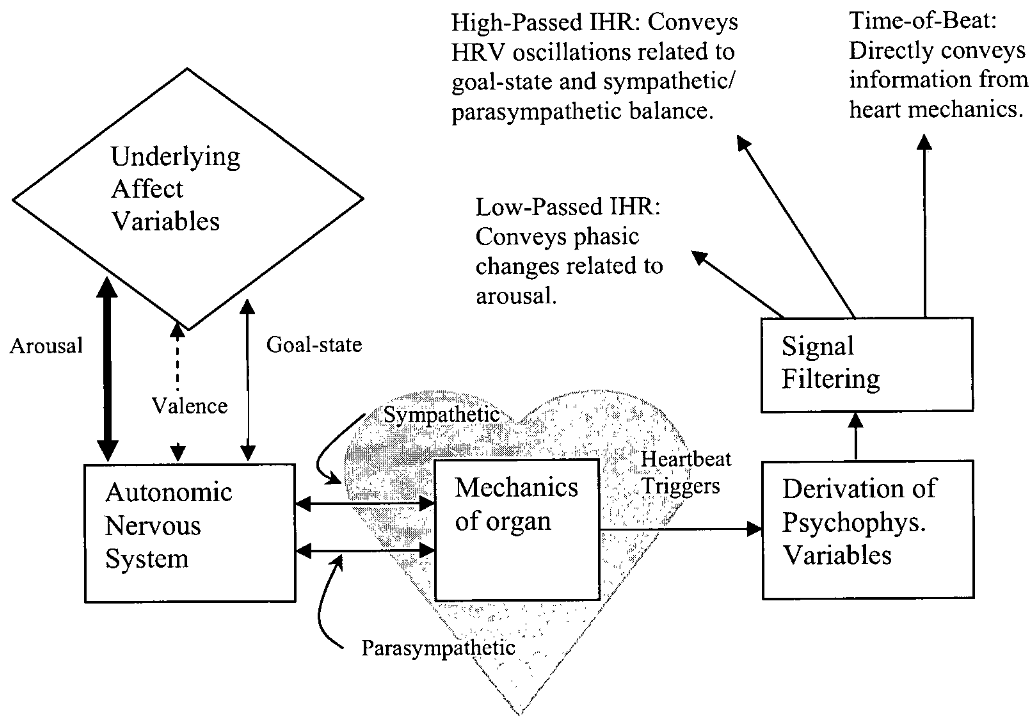


Figure 10-2. Design Approach to Processing IHR Variables in the 2Hearts System. The system outputs relate most strongly to arousal, ANS balance and heart mechanics.

Specific Inter-User HR Variables Chosen

The co-navigation paradigm includes pathways providing information to the users about inter-user heartbeat patterns. The derivation and interpretation of these variable is more speculative, since a clear natural model for the production and perception of these variables has not been identified.

The inter-user heartbeat measures chosen from the possibilities in Section 5.5 were correlation(IHR1,IHR2), difference(IHR1,IHR2) and mean(IHR1,IHR2). The correlation measure is an interesting hypothesized test for relative frequency synchronization, while the difference measure addresses the phenomena of relative frequency synchronization in a direct/low-level way. Mean(IHR1,IHR2) was included to portray overall/common arousal state between users. The display of time-of-beat information also provides the opportunity for absolute/relative phase synchronization phenomena to be observed.

In summary, four major themes define the psychophysiological design of 2Hearts:

1. a focus on the dimension of arousal (and to a lesser extent, goal-state), due to the use of the heart as sole visceral data source
2. initial constraint to heartbeat variables measurable from the Polar sensor (i.e. IBI-derived measures)
3. a preference for exploring the unconscious representations of affect rather than conscious/categorized representations of affect
4. the presentation of HR/HRV patterns in a fairly 'raw' (not highly processed) form, as a starting point and to investigate the direct perception of such rhythmic forms

Pre-Processing Sensor Data

Primary considerations for designing SBFS sensor pre-processing methods are rejecting noise, capturing the necessary physiological data of interest, and creating a dynamic, realtime, low-latency processing stream to maintain affective contour information. The Polar sensors use an analog circuit to generate a trigger pulse once per detected QRS complex. The trigger pulse was observed to maintain a similar shape on every beat for normal heartbeat intervals. To obtain precise time-of-beat readings for digital processing, it was decided to digitize the trigger pulse waveform and use peak detection on the resulting data to calculate beat-onset times. It was planned to implement a simple algorithm to filter out/remove beats judged as too closely spaced to be physiologically possible.

10.3 Design Phase III: System Outputs

This section lists the many design parameters chosen for the 2Hearts—Version 3 audio-display. Many of the design decisions involved weighing benefits and drawbacks of potential audio parameters- the final chosen set of parameters is explained, with benefits/drawbacks pointed out. Note: For specific values of constants used in this design, see Section 10.5.

Filtering

To aid perception of various phasic changes and pattern phenomena in the IHR signal, a filtering scheme was designed to split each user's IHR data into two display streams: one lowpass filtered stream and one highpass filtered stream. The splitting strategy was used partly in response to the observation from early prototypes that the regular up-and-down pattern of breathing-linked HRV tended to dominate perception and obscure other phasic HR changes.

The cutoff frequency of the lowpass stream was chosen through informal testing with the goal of damping oscillitory HRV phenomena, particularly those related to respiration, leaving apparent medium-term phasic HR changes such as periods of excitement or stress. The cutoff frequency of the highpass stream was adjusted to highlight patterns of HR variation including HRV oscillations, orientation and expectation responses. Filtering was also used on the difference(IHR1,IHR2) signal to focus attention on phasic changes rather than the constant difference in tonic level between two given users. It will be explained below that the employed filtering scheme requires further development to function as intended.

Time Scaling/Looping

Looping was not employed for the user 1, 2 HR measures since as discussed in Section 6.1 it is possible that a natural bias will lead to these patterns being optimally recognized when displayed at their original time scales. However, the inter-user heartbeat measures are less likely to be directly linked to internal contour representations of affect. Therefore a looping scheme was designed for one of the relative frequency synchronization measures to explore whether inter-user heart phenomena could be detected in sped-up data. The difference(IHR1,IHR2) stream was time-compressed by a factor of roughly 5:1 (see Section 10.5), indicated by informal observation to map the most obvious difference patterns into the perceptual timescale of "rhythm".

Audio Stream Grouping

As indicated in the SBFS design cycle, the principle of stream separation was employed to aid users in understanding and being correctly influenced by the audio modulation parameters.

To assist users in attributing the several per-user heart rate variables (time-of-beat 1,2, IHR 1,2 lowpassed/highpassed) to a common source, these variables were assigned to a single perceptual stream. In this stream, modulation values are updated once per beat, with the time-of-beat measure directly triggering the onset of a new stream amplitude envelope. This discretization in time does not discard information since IHR can only be measured once per heart beat. To allow comparability between parallel psychophysiological measures of the two users, two separate streams with these same mappings were created, one for each user, distinguished by timbre signature and stereo pan position. To distinguish and group the inter-user HR measures and create the metaphor of a 'third entity', formed by the users' social interrelation, a third stream was assigned. This third correlation stream was given similar parameter design to the first two.

Hence, a three-stream concept was developed which could be informally phrased as "me, you , and us". The concept was later expanded to four streams in order to incorporate separately the correlation(IHR1,IHR2) and difference(IHR1,IHR2) measures. Keeping in mind the inter-display-channel possibility of interference, the timbral frequency distribution of each stream was designed to minimize overlap in frequency space, as discussed below in the individual stream sound design descriptions. A 'drum beat' stream was also added, driven by mean(IHR1,IHR2), to encourage perception of rhythms in the data as being 'musical'.

Global Sound Design

One global sound design parameter affecting every audio modulation parameter is the total audible volume of the mixed audio stream. Global volume affects the relative intensity and clarity of the user's vocal sounds versus the system audio. The system was designed to be adjusted to a comfortable listening volume similar to but slightly louder than typical background music.

Sound Design/Bias Adjustment: 'Lead' Streams for Heartrates 1,2

Each of the two per-user HR-linked streams was given a similar but distinct timbre design. One stream was given a synthetic timbre reminiscent of a string section, the other given a timbre reminiscent of a horn section. As an unassigned audio parameter, amplitude envelope shape was chosen to be distinct but not overly sharp, in order to maintain associations with a discrete 'beat' event but minimize associations of high-arousal or hard physical impact. Stereo pan position for the two streams was set to hard left and hard right, respectively, to further distinguish between per-user signals. The timbre design and typical pitch of these streams combined to create a spectrum concentrated in the mid/high-mid audio range. This range is distinct from other 2Hearts audio streams, and maximizes accuracy in pitch and melody perception. This range may carry unwanted associations of high arousal, and tend to draw the focus of conscious attention.

Selection of Audio Modulation Parameters: 'Lead' Streams

The following lists describe the parameter modulation scheme that was created for the per-user HR –linked audio stream:

<u>Mapping: Time-Of-Beat → Note Onset Time</u>

- **Explanation of Mapping:** Each new beat triggers onset of an amplitude envelope for a new note in the 'lead' stream (the previous amplitude envelope is cut short if necessary). The steady nature and time-scale of heartbeats leads to perception of the high-level variable of tempo, in parallel with perception of rhythmic offsets in timing. This mapping also allows the perception of inter-user heart beat phase synch. The direct synchronization with one user's heartbeat may encourage user identification with the signal.
- **Inherent Bias:** Onset of a discrete note carries the correct associations of a discrete physical event (heartbeat). Increasing tempo of notes carries the correct associations of increasing speed and arousal.
- **Estimated Display-Channel Characteristics:** Tempo has good long-term memorability. Perceptual accuracy of small changes in tempo is medium.
- **Range Remapping/Warping:** Not used.

Mapping: HR 1,2 → Lowpass Filter → Stream Pitch

- **Explanation of Mapping:** The filtered HR value (updated once per beat) is used to modulate pitch of the 'lead' stream.
- **Inherent Bias:** Higher pitch carries the correct associations of higher arousal and faster speed. Other undesired associations may result from random dissonance between serial notes and between the two pitch streams. The up/down pattern of pitch created by a typical HR signal may convey additional unknown associations.
- **Estimated Display-Channel Characteristics:** Accuracy of pitch perception is estimated to be greater then accuracy of tempo deviation for displaying the same data according to the 2Hearts system mapping. If pitch patterns are perceived as melody, memorability and comparability of patterns will be enhanced, although unwanted bias may result. Inter-channel comparability between the two pitch streams should be high.
- **Range Remapping/Warping:** A linear mapping from (lowpassed) HR in BPM to pitch in Hz was used.

Mapping: HR 1,2 → Highpass Filter → Stream Brightness

- **Explanation of Mapping:** The filtered HR value (updated once per beat) is used to modulate filter-cutoff of a low-pass filter applied to the 'lead' stream.
- **Inherent Bias:** Higher brightness carries the correct association of higher arousal. It may carry unwanted associations of unpleasantness. Cyclically fluctuating brightness may convey additional unknown associations.
- **Estimated Display-Channel Characteristics:** Long-term memory for absolute values may be poor. Accuracy of perceiving moment-to-moment changes in brightness may be medium to good; patterns will be perceived in terms of rhythm. If distinctive patterns are present, memorability for these rhythms may be high.

- **Range Remapping/Warping:** A linear mapping from (highpassed) HR to filter cutoff-frequency in Hz is used.

Sound Design/Bias Adjustment: 'Bassline' Stream for HR-HR Correlation

To form a non-interfering counterpoint to the per-user HR streams, the joint user-user correlation stream was designed as a lower-frequency 'bassline'.
The sound design chosen was a simple subtractive-synthesis bass: a sawtooth oscillator through an audio filter. (Simple synthetic sounds are often more aesthetically pleasing at low frequency ranges.) It was discovered in early testing that the bassline needed another source of variation in order to sound natural and 'musical', hence, a subtle arbitrary volume modulation pattern was added to the bassline (see Section 10.5 below for details), in synchronization with the loop frequency of drum and noise streams (essentially a synchronized 'rhythm section' based on users' joint HR behaviour).

Selection of Audio Modulation Parameters: 'Bassline' Stream

The following lists describe the parameter modulation scheme that was created for the user-user HR correlation –linked audio stream:

<u>Mapping: Mean(HR1,HR2)→Stream Pitch</u>

- **Explanation of Mapping**: The mean value of users' heart rates modulates the pitch of the 'bassline' stream. This mapping parallels the 'lead' stream, aiding the impressing of the user-user correlation as a 'third person'
- **Inherent Bias:** Higher pitch conveys the desired association of higher arousal. Higher pitch in bass range is less likely to create dissonance leading to negative valence.

- **Estimated Display-Channel Characteristics**: Accuracy/Memorability in the bass range may be lower than treble range, but still good. Relative perception of pitch contours should be fairly accurate.
- **Range Remapping/Warping:** A linear mapping is used from Mean HR in BPM to pitch in Hz.

Mapping: Correlation(HR1, HR2)→Stream Brightness

- **Explanation of Mapping**: The correlation between users' heart rates modulates the filter cutoff of the 'bassline' stream. This mapping attempts to portray episodes of low user-user correlation as representing social friction; ie. High arousal and negative valence.
- **Inherent Bias:** Higher brightness conveys the desired association of arousal. Higher brightness may also convey negative valence, which for this mapping would be appropriate.
- **Estimated Display-Channel Characteristics**: Accuracy of brightness perception in bass range will be quite low; but rough level and extreme changes should be obvious.
- **Range Remapping/Warping:** An inverse mapping is used between correlation and stream brightness.

Sound Design/Bias Adjustment: 'Noise' Stream for HR-HR Difference

As a parallel/alternate attempt to render perceivable joint user-user aspects of HR behaviour, the difference in user HR's was mapped to a high-frequency filtered-noise stream. The noise stream was intended to occupy a higher frequency range then the lead lines, preventing interference between streams. The noise stream was intended to complement the 'drum' and 'bassline' streams. Timbre design was created with a noise oscillator run through a bandpass filter. Additional fixed audio filtering was done by trial and error to give the noise a more neutral, non-harsh character.

Selection of Audio Modulation Parameters: 'Noise' Stream

The following lists describe the parameter modulation scheme that was created for the user-user HR difference-linked audio stream:

Mapping: Time-compressed Difference(HR1, HR2)→Stream Centre Frequency

- **Explanation of Mapping**: A sliding/looping window of HR-HR difference values controls the centre-frequency of the bandpassed noise stream
- **Inherent Bias:** Relative centre-frequency increase should convey increased arousal and possibly negative valence; both appropriate for increase user-user difference. The patterns of variation in frequency may also convey unknown associations.
- **Estimated Display-Channel Characteristics**: Accuracy of centre-frequency perception is probably low; but looping pattern behaviour should be perceived as memorable rhythms, thus exploring potential inter-user heartbeat relationships
- **Range Remapping/Warping:** A logarithmic mapping from HR difference to filter frequency was used (it was found to give a more 'linear sounding' mapping than a linear mapping; not surprisingly given human perception of frequency)

Mapping: Time-compressed Difference(HR1, HR2)→Lowpass Filter→Stream Amplitude

- **Explanation of Mapping**: To make the noise stream more noticeable during episodes of high HR difference, a smoothed difference signal controls the volume of the stream
- **Inherent Bias:** Increase volume of noise should convey increased arousal, a desired bias, and may also convey negative valence, which would be suitable.
- **Estimated Display-Channel Characteristics**: Amplitude should have low accuracy of perception, which is acceptable given the goal of simply highlighting certain episodes.

- **Range Remapping/Warping:** Stream amplitude varies according to the squared value of HR difference in order to highlight extreme values of difference.

Mapping: Mean(HR1, HR2)→Stream Looping Tempo

- **Explanation of Mapping**: To synchronize the time-compressed rhythms into a musically sensible 'rhythm section', noise looping speed is controlled by mean heart rate.
- **Inherent Bias:** Faster rhythms convey increase arousal, which is appropriate; faster noise variations will also increase the perception of 'roughness/fluctuation'
- **Estimated Display-Channel Characteristics**: Long-term memory of tempo is good. Perceptual accuracy for changes in tempo is medium.
- **Range Remapping/Warping:** Not used.

Sound Design/Bias Adjustment: 'Drumbeat' Stream for HR-HR Mean

As mentioned above, a final 'drumbeat' audio stream was employed mainly for the function of tying together perception of the various audio rhythms into a gestalt more likely to be perceived as music. The timbre of the drumbeat was taken from an audio sample of a typical 'dance' style kick drum, in order to emphasize associations with popular music.

Selection of Audio Modulation Parameters: 'Drumbeat' Stream

Mapping: Mean(HR1, HR2)→Stream Looping Tempo

- **Explanation of Mapping**: The tempo of users' mean HR mapped directly into the drum beat tempo.
- **Inherent Bias:** Faster drumbeats convey higher arousal, a desired bias.

- **Estimated Display-Channel Characteristics**: Long-term memory of tempo is good. Perceptual accuracy for changes in tempo is medium.
- **Range Remapping/Warping:** Not used.

Summary of Mappings

Table 10-1 lists the designed streams and mappings. Desired relative bias associations are listed; the presence of these desired associations was tested empirically as described in Section 12.2.

Stream Description	Input	Output	Desired Relative Bias
'Lead' Pitch Sequences	Time-of-beat	Note onset time	Faster = higher arousal
	HR (thru LPF)	Pitch	Higher pitch = higher arousal
	HR (thru HPF)	Brightness	Brighter = higher arousal
'Bassline'	Mean HR	Pitch	Higher pitch = higher arousal
	HR-HR Correlation	Brightness	Brighter = higher arousal, brighter = negative valence
'Noise'	HR-HR Difference	Centre Frequency	Brighter = higher arousal, brighter = negative valence
	HR-HR Diff (thru LPF)	Amplitude	Louder = higher arousal, louder = negative valence
	Mean HR	Looping Tempo	Faster = higher arousal
'Drumbeat'	Mean HR	Looping Tempo	Faster beat = higher arousal

Table 10-1. Selected Aspects of 2Hearts--Version 3 Audio Display Mapping.

10.4 Hardware Implementation

Heart Sensors

The 2Hearts—Version 3 system uses inexpensive (~$90 Cdn) Polar heart monitors resold by the Vernier company as model EHR-DIN [85] [87]. Each monitor package includes one chest-strap/transmitter and one standalone receiver module with a DIN connector supplying the raw received trigger pulse signal.

Certain more expensive models of Polar heart sensor employ coded transmission to avoid cross-sensor interference. It has not been determined whether standalone receivers capable of supplying the raw trigger signal are available for such models. For the present work, a modification was undertaken of one of the inexpensive sensors to prevent interference: physical distancing of the transmission module from the chest electrodes by cutting and extending the electrode/transmitter connection. In this case, new electrodes were created using thin, paper-backed copper sheets available at a craft shop.

Figure 10-3 shows the modified Polar transmitter. (A knife was used to cut away the soft chest-strap plastic and reveal the transmitter's input connections.) It was found by trial-and error that the added 10kΩ resistors compensated for the lower resistance of the copper sheeting compared to the original conductive-plastic electrodes, allowing the transmitter to function normally.

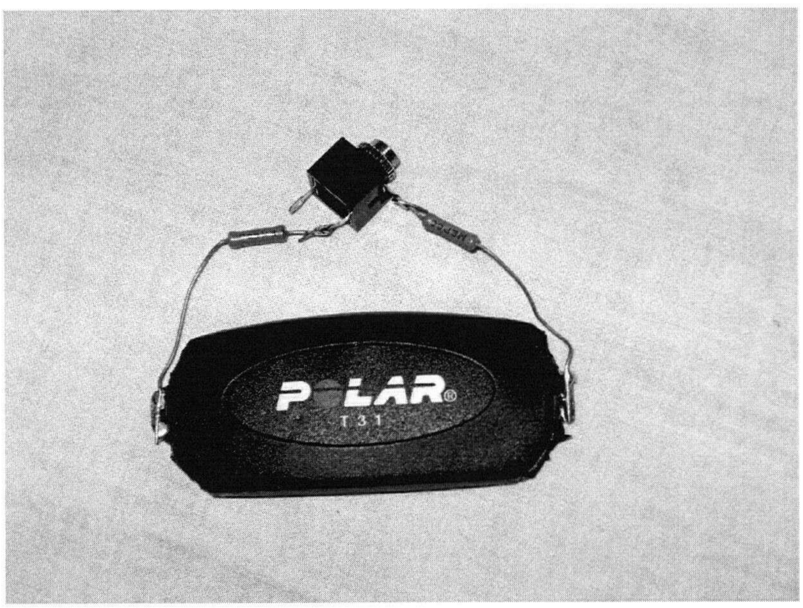


Figure 10-3. Modified Heart Sensor (Transmitter). Note the stereo minijack attached for extending the inputs.

Figure 10-4 shows the replacement chest belt that was created by attaching rectangles of copper with two sided tape to a padded foam strap. (This configuration fit together with the removable behind-back strap from the original Polar chest belt.)

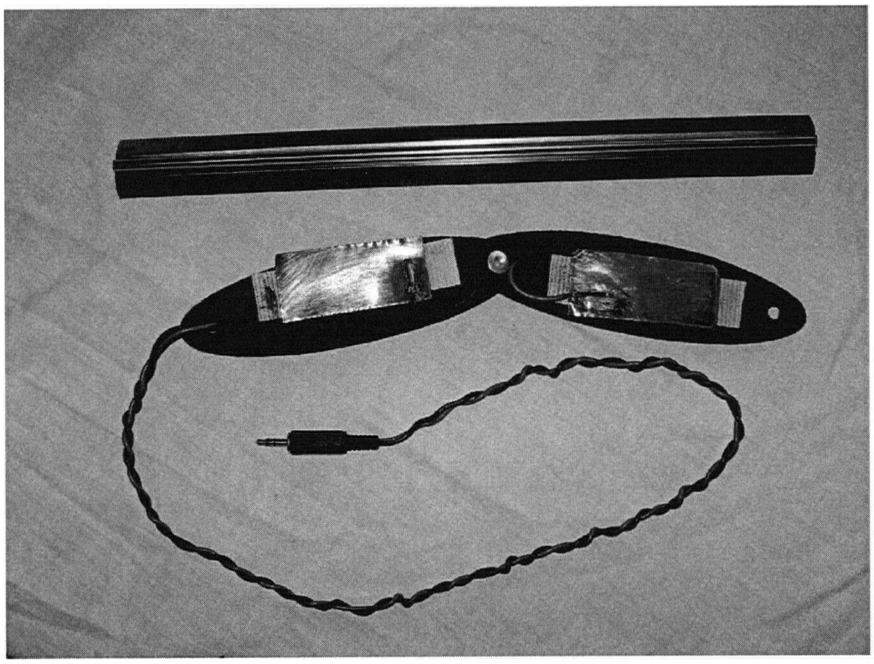


Figure 10-4. Modified Heart Sensor (Electrodes). Note the stereo mini-plug attached form extending the outputs.

Heart Signal Digitization

After trying other methods of digitizing heart-sensor data (see Section 8.2), it was realized that a standard but high-quality PC audio interface is capable of digitizing the required trigger signals with high sample-rate/bit-depth, low jitter, and low latency; and supplying the data stream in a format readable natively by most audio software.

Therefore, the A/D converter used in 2Hearts—Version 3 implementation is the Echo Indigo I/O PCMCIA card, running at 16 bit/44.1kHz. Although the card is AC-coupled (and therefore not designed to support arbitrary control signals), visual inspection of the digitized triggers showed that consistent-shaped, sharp, peaks were reliably present.

Processing and Display Hardware

The PC used to run 2Hearts software is a Pentium 4, 2GHz Dell Inspiron 4150 Laptop with 512MB of RAM. The audio amplifier used for sound output is a Yamaha AX-592 and the speakers are Paradigm Atoms.

System Diagram

Figure 10-5 below shows the components of the 2Hearts--Version 3 system.

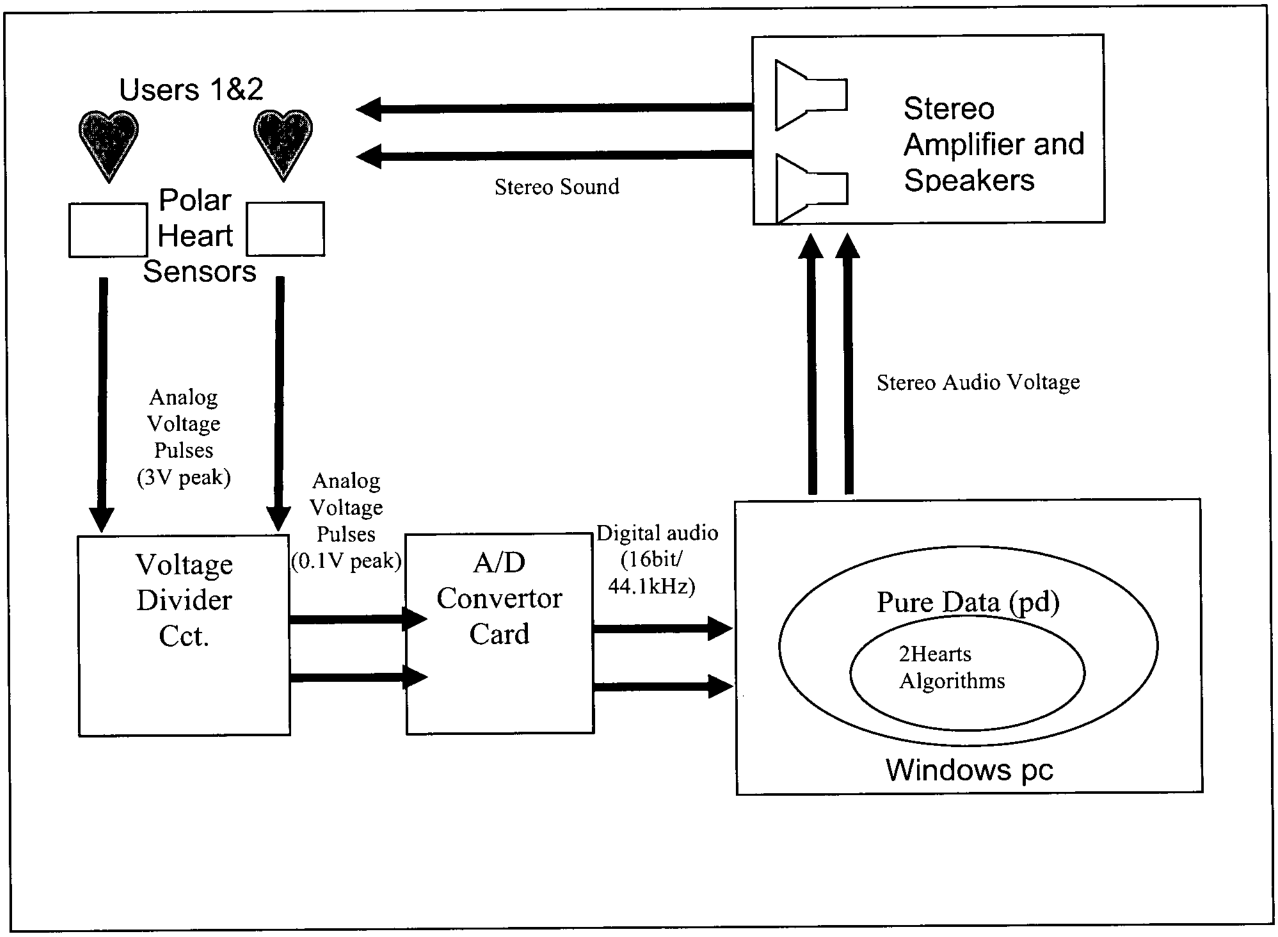


Figure 10-5. Components of the 2Hearts--Version 3 Implementation. The test system also employed video feedback using an LCD projector (not shown) as a control condition (not considered part of the system proper.)

10.5 Software Implementation

The 2Hearts—Version 3 software is implemented within the Pure Data audio-signal-processing environment, running under Windows XP. The 2Hearts software is stored as a Pure Data 'patch' file representing a configuration of connected low-level signal-processing operations, structures and parameters. The low-level PD objects are grouped into functional 'subpatches' which take inputs and outputs that represent audio signal flow or control signal flow. Figure 10-6 shows the main PD window, displaying a high-level view of the 2Hearts--Version 3 patch.

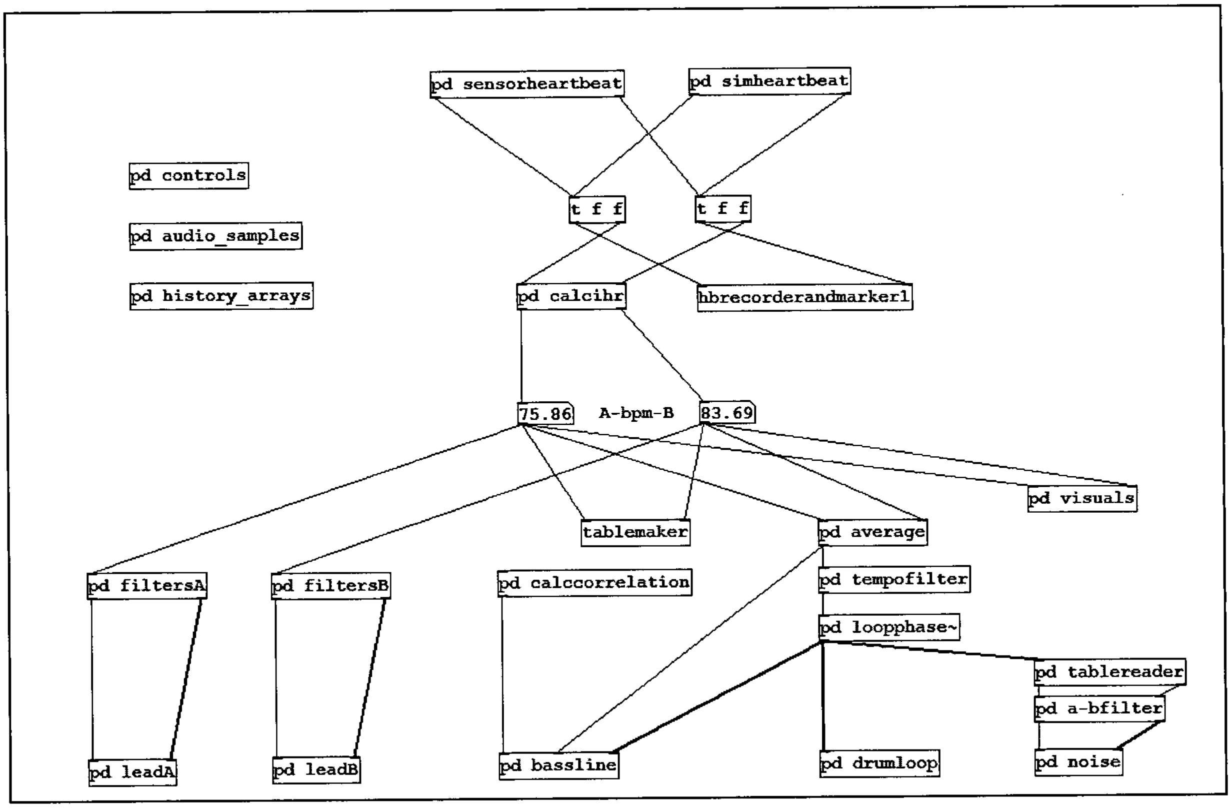


Figure 10-6. Main PD Patch Window of 2Hearts--Version 3. Audio output modules are grouped along the bottom row, and heart data input sources along the top. The boxes containing "t f f" are PD 'trigger' routing objects which pass incoming data to both outputs, in right-output then left-output order. Note: Stored sample and array data can be accessed by named reference within the PD environment without the need for explicit signal connections.

A notable aspect of the system software architecture is the use of 2 circular arrays to store a continuously updated window of each user's past HR data. The 'tablemaker' object writes interpolated data into this array at a constant rate. The 'calccorrelation' object reads stored array data to calculate inter-user IHR correlation for the windowed time interval. The 'tablereader' object traverses the array repeatedly at a varying rate depending on mean(IHR1, IHR2) to produce a looping/sliding window used in the display of difference(IHR1, IHR2).

To characterize this looping scheme numerically: PD is configured to run at a sample rate of 44.1kHz, with audio processed in digital-signal-processing (DSP) 'chunks' of 64 samples. Processing time for one 'chunk' is called a 'tick'. The driver latency of the A/D input card is set to 128 samples. Heart rate data is sampled into the arrays once per 10 DSP ticks, and the arrays are 1000 samples in length. Therefore each array holds about 14.5 seconds of HR data. The 'tablereader' object traverses the array once for every 4 beats of a virtual 'mean heart rate'; thus for a mean heart rate of 80 BPM array data is read out as time-compressed by a ratio of about 4.84:1.

Appendix A details the operation of the 2Hearts-- Version 3 software by describing the inputs/outputs and functionality of each subpatch in turn, and Appendix B contains a code listing for the custom-coded portion of the software.

11. Design of 2Hearts User Testing

The 2Hearts user testing was designed and performed in conjunction with the development of the SBFS design cycle (see Section 7.3). Therefore the user testing did not correspond exactly with one of the suggested iterative testing phases, but rather was an exploratory set of tests that incorporated elements of several testing phases. The experience of 2Hearts user testing was then used to inform the SBFS design cycle.

11.1 Objectives, Task/Questionnaire Design and Predictions

This section describes the selection and specification of the different tasks used to test aspects of the 2Hearts system.

Objectives

As a relatively new field, SBFS design does not have well-established test paradigms or system performance measures. One high-level objective of the 2Hearts user testing was to explore several different types of measurement technique and stimuli/ social context design to evaluate their usefulness for SBFS design.

Other, more specific, objectives for the 2Hearts user testing included:

1. To evaluate the system's input/display mapping design by examining both conscious and unconscious effects with respect to bias and other DCCs, and gain ideas about how to improve the mapping
2. To measure the total system performance of the current design, with respect to conscious/unconscious characteristics, in the social context of the 'co-navigate' SBFS paradigm and other subsets of this paradigm

3. To obtain qualitative user opinions/impressions of the overall system-use experience and suitability of such a system for various applications
4. To investigate the use of between-user HR cross-correlation as a measure of social connectedness or empathy
5. To acquire heartbeat data to use in future system design (as sample data for testing input/display schemes)
6. To acquire sample conversational audio/video linked to heartbeat data for use in future experimental design and theoretical development (enabling independent conversational measures of empathy, or social-event triggers of heartbeat behaviour, to be explored)
7. To acquire baseline sample data regarding resting HR and the overall effect of system audio on HR; thus identifying a potential global bias of system audio on HR.
8. To gain understanding of how large/significant are typical physiological effects of audio feedback, in order to aid future theory/experiment development

Design and Rationale for Tasks I-III

One strategy used to address the above objectives was to measure user response to the system by measuring user self-reported and rated perceptions, and by measuring the corresponding heart signals. Self-report was intended to measure the result of the pathway S→C, (and perhaps S→U→C and S→U→V→C), while heart measurement was expected to provide more focus on the results of pathway S→U→V (and hence insight into S→U).

Three tasks were invented to allow separate testing of 2Hearts input/display mapping and testing of system performance in different social contexts; the tasks are referred to as Tasks I to III. Briefly, the tasks were:

I. The baseline/listening task, in which subjects did not interact with each other, and listened to musical excerpts or silence.

II. The conversation task, in which subjects conversed with each other while listening to real, spoofed, or silenced feedback from the 2Hearts system.

III. The control-game task, in which subjects took turns trying to deliberately influence each others' heart rates while receiving audio/video feedback or video-only feedback.

Figure 11-1 shows a visual summary of each Task. The Tasks are described further below.

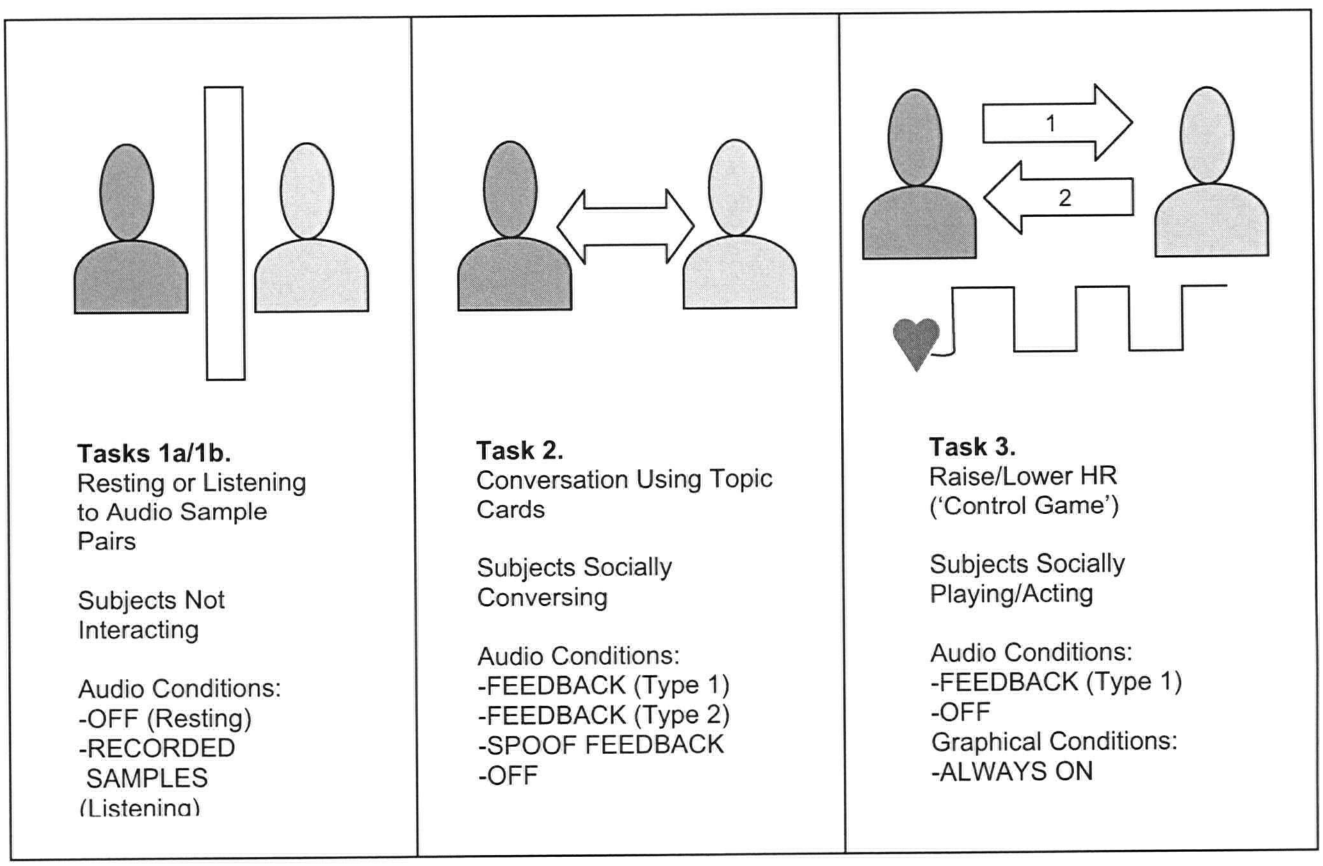


Figure 11-1. Graphic Depiction of 2Hearts User Testing Tasks I-III.

Description/Rationale for Task I

Task I was intended to evaluate the directional bias DCC for certain audio modulation parameters employed in the 2Hearts system. Four pairs of one-minute audio excerpts were recorded in advance from 2Hearts output, by setting various constants within the audio-mapping algorithms to different values per excerpt, and muting different audio streams to isolate the stream including the parameters of interest. (Given this isolation, inter-stream interference/comparison was not directly explored.) The amount of parameter-change within each pair was chosen by estimation, based on previous informal system testing, to simulate two contrasting values that might occur under normal system use. (This specification of audio behaviour, based on parameters of the 2Hearts system, was intended as a heuristic method of use in system design/ prototyping, rather than a method for generally-applicable psychoacoustic research.)

The four audio pairs were as follows:

1. Dual/parallel pitch sequences (Audio parameter changed = mean pitch): This pair isolated the two 2Hearts per-user heartbeat/pitch sequences. The scaling constant of the HR-to-pitch mapping was set to either 6 or 8 (ie. a 33% increase in mean pitch frequency). The sequences still contained random up/down pitch changes from note to note, controlled by the 2Hearts random input-simulation module (with one simulated stream set to create random inter-beat periods between 600ms and 820ms and the other set for periods of 620ms to 790ms - See Section 10.5 for explanation).

2. Single series of percussive tones (Audio parameter changed = tempo): This pair isolated the 2Hearts mean HR/drumbeat stream. Each pair member was an evenly-spaced series of percussive beat sounds, one at 75 BPM and one at 100 BPM.

3. Continuous low-pass filtered bass tone with modulated brightness (Audio parameter changed = mean brightness): This pair isolated the bass tone used for display of HR-HR correlation data in the 2Hearts system. The additive mapping constant for filter brightness was varied to give one pair member a mean filter-cutoff of approx. 450 Hz higher than the other. Random variation in the brightness was created by employing the 2Hearts random input-simulation module with values set as in audio pair 1 above.

4. Continuous bandpassed white noise with modulated bandpass-centre-frequency (Audio parameter changed = roughness of dynamic variation in filter-centre-frequency): This pair isolated the noise sound used for display of HR-HR difference data in the 2Hearts system. The smoothing filter cutoff-constant for the difference data was varied between 0.25Hz and 1.5Hz. Random variation in the centre-frequency was created by employing the 2Hearts random input-simulation module with values set as in audio pair 1 above.

The fourth audio pair was designed to explore the perceptual bias relating to the perception of dynamically-varying timbre modulation versus more smoothly varying timbre modulation; this distinction was not an explicitly modulated parameter in 2Hearts but rather an effect that may occur in the perception of different heartrate patterns.

A resting-in-silence phase was included to gather baseline data for different subjects' HR and HRV, and to help measure whether system audio caused a measurable offset in HR.

Description/Rationale for Task II

The ultimate design goal of the 2Hearts system relates to the complex *co-navigate* SBFS paradigm. It was reasoned that more tractable testing strategies for initial testing might be found by decomposing the *co-navigate* paradigm into the simpler paradigms of *control, understand, and converge* (See Section 3.3).

Task II was intended to test aspects of 2Hearts performance relating to the *understand* and *converge* paradigms, in a social context where the system is used as 'background' to augment a naturalistic, social conversation. In terms of the signal flow model, Task II was intended to test U1→U2 and U1→C2 transmission. Transmission to C was intended to be measured by questionnaire/self-report, and transmission to U was intended to be measured by heart rate behaviour.

Given the wide variation in mood/pace/content social conversations, it was desirable to find a test paradigm that fixes conversation within a standardized framework. A commercial board game, *Pharoah's Bluff*, was identified, to supply useful 'conversation topic cards'. The cards list sets of 4 questions designed to initiate meaningful conversations.

Since this system-usage context employs system feedback alongside an already rich set of social communication signals, it was expected that perceptual bandwidth/attention would be relevant concerns. Thus, two versions of 2Hearts input/display mappings were tested, the full mapping (see Chapter 9) and a simplified mapping lacking the difference- and correlation- based audio streams.

To discriminate the effects of actual heart-data feedback versus the raw effects of system audio, a spoof condition was designed. The spoof data employed the full audio display and was based on actual per-user heart data, delayed by 2 minutes. This method was intended to prevent users from recognizing or reacting differently to the spoof condition due to obvious difference in overall HR mean/variation. It was reasoned that a time-delay of 2 minutes would sufficiently decorrelate the audio feedback from the actual conversational behaviour. A no-audio condition was also employed.

Thus, the 4 system-audio test conditions were silence, spoof feedback, real feedback, and real feedback lacking correlation signals. The order of the 4 audio conditions was set differently for each of the 4 subject pairs, using a latin square design.

Description/Rationale for Task III

Task III was intended to test whether the 2Hearts SBFS functioned well according to the goals of the SBFS 'control' paradigm (see Section 3.3). In terms of the signal flow model, it was intended to test whether the system action led to enhanced intentional control of user 2's heart rate by enhancing pathways U2→C1 and U2→U2. (No pathway damping action was included.) Task III was also intended to test the more general question of whether a person is capable of purposefully controlling the heart rate of another using social-interaction methods. Video feedback was included as the control condition,

For a given task segment, one user was instructed to influence the other's heart rate in a specified direction, using arbitrary communicative strategies but not physical contact. The up/down goals were to alternate in order to simplify the consideration of ordering effects. The agreement of measured HR behaviour of the target user with goal function was chosen as the measure of success of the overall paradigm.

Each subject's heart rate was displayed by the 2Hearts system according to the simplified mapping used in Task II (see above). Video feedback was also employed by projecting the 2Hearts visual bar-graph (see Section 10.5) on a wall visible to subjects. Task segments using video-feedback only were designed to compare with task segments using audio and video feedback, in order to show whether the complex audio mapping design presented any advantage over a simple informational video mapping.

Thus, the 3 independent variables for the task were **system audio** = [on, off], **target user** = [subject 1, subject 2], and **HR direction goal** = [up, down]. Using the constraint that the HR direction goal always alternated, these 3 variables were used in different combinations for a total of 8 segments. Each segment lasted for 2 minutes. The order of combinations was partially counterbalanced between groups. The resulting data set was not fully balanced since the data from one subject pair was discarded for this task (the task was not completed due to a subject's scheduling needs.)

Design of Questionnaire

A questionnaire was designed to gather user self-report information from Tasks I and II, and to gather additional information about users' opinions regarding the enjoyability, effects, and speculated utility of the 2Hearts system (see Appendix E for a copy of the Questionnaire).

The questionnaire had three parts:

I. Rating and free-form responses related to each audio-pair excerpt.

II. Rating and free-form responses related to each phase of the conversation task.

III. Rating and free-form responses related to overall experiences with the system. (Note that Questionnaire Part III does not correspond directly with Task III).

Brief explanations of the three parts follow:

<u>Questionnaire Part I</u>

For each pair of audio excerpts, subjects had to compare excerpts in terms of which sounded 'more energetic' by making a vertical mark on a continuous horizontal scale. A second, similar scale was made for the comparison of 'happier'. The midpoint of the scale was indicated as 'neutral'. The terms 'happier'/ 'more energetic' were chosen as layperson-equivalents to the concepts of valence and arousal. The scales were designed to simplify the analysis of within-audio-pair comparisons by obtaining ratings directly in terms of within-audio-pair differences. By using a continuous scale, it was speculated that subjects could respond in a more direct/instinctive way. (In hindsight, the desire for instinctive response may conflict with the use of questionnaires to measure 'conscious' response; at the time the questionnaire was created, the theoretical framework was not in place to clarify this issue.)

Two freeform questions asked subjects to contrast, within an excerpt pair, how each excerpt made the subject feel, and how the excerpt sounded. These questions were intended to gather qualitative data about the affective/metaphorical associations of each excerpt, and about subject's intellectual understanding of the nature of the given audio-parameter variation.

Questionnaire Part II

Subjects were given 6 questions regarding their experience of system audio. For each question, 4 continuous horizontal scales were given to comparatively rate the 4 audio-condition-segments. This section was intended to discover if subjects consciously noticed a difference in the output or social impact of the system with respect to the different audio-feedback conditions.

Two freeform sections asked subjects, for each conversational segment, to describe their conversational experience and their observations regarding system audio output. These sections were intended to probe the nature of influences the 2Hearts system's action might exert on the experience of conversation, discover what types of information subjects were consciously recognizing in system output, and further explore the affective connotations of system audio when experienced in the social context.

Questionnaire Part III

Questionnaire Part III included 5 statements about the overall system-use experience, to be rated on a 5-point disagree/agree scale. It also include 3 freeform-response questions, about the system sounds, enjoyability of the system, and speculated applications of a 2Hearts-style system. Miscellaneous questions about subject's caffeine use, musical training, and before-hand familiarity with the other subject were included; these topics were thought to be potentially useful for analyzing/explaining marked deviations in a particular subject or pair's data.

Predictions

The following hypotheses were made in advance of user testing as guidelines for data analysis.

H1. Subject's mean heart rates will show an overall within-subject increase for the listening phase as compared to the resting phase. (Null hypothesis: The population mean of the per-subject difference in mean heart rates between resting and listening is equal to zero.) This hypothesis was made due to test whether the suggested tendency of music to increase heart rate, as discussed in Section 6.3, was found for the 2Hearts system audio.

H2. Given paired samples of musical output from the system, subjects will rate paired samples, within each pair, along dimension of "more energetic/less energetic" and "more happy/less happy" in agreement with the designed parameter mappings. (Null hypothesis: the population mean of the per-subject difference in ratings, for a given audio-sample pair, is equal to zero.)

H3. During the conversation task, the zero-lag cross-correlation of the user-user heart rate signal will be greater for the experimental conditions of <system is enabled- type 1> and

<system is enabled- type 2>, than for the conditions of <system is disabled> or <system is using delayed input data>. (Null hypothesis: There is no difference in population means of heart-correlation for the different system conditions.)

H4. Rating their experiences during the conversation task in several categories, subjects will rate the type1 and type2 phase differently from the spoof (delayed) phase. (Null hypothesis: the different phase conditions have no effect on subject ratings.)

H5. During the relax/excite task, the target user's mean heart rate per segment will show an above-chance correlation with the independent variable **HR direction goal**. (Null hypothesis: the value of **HR direction goal** has no effect on targer user's mean heart rate.)

H6. During the relax/excite task, the target user's mean heart rate per task segment will show a higher correlation with the independent variable **HR direction goal** during the experimental condition of <system is enabled>, as compared to the conditions of <system is disabled>. (Null hypothesis: the effect of **HR direction goal** does not depend on the condition of **system is enabled/disabled.**)

Once the data were gathered and examined, several other potential trends in the data were also analyzed (see Section 12.2).

11.2 User Testing Methodology

This section describes the specific equipment and methods used to carry out the 2Hearts—Version 3 user testing.

Subject Selection

Subjects consisted of members of the University of British Columbia's Human Communication Technologies Laboratory (HCT Lab) and the friends of lab members. Subjects were contacted via word-of-mouth and via a poster on the door of the HCT Lab. Potential subjects were informed briefly that the experiment involved listening and socializing tasks with heart-rate monitors and audio feedback, then directed to a web page for further information. The web page contained information about the experiment, listed eligibility requirements, and contact information for volunteering. To be eligible, subjects could not have pacemakers or known medical heart problems. All subjects were between 22 and 31 years old. Eight subjects were used, 5 male and 3 female. Subjects were arbitrarily assigned to pairs on the basis of scheduling constraints, resulting in 3 male/female pairs and 1 male/male pair.

List of Apparatus

- 2Hearts system hardware/software as listed in Section 10.4 including heart sensors, laptop, amplifier and speakers
- Sony mini-DV digital video camera
- Tripod for the video camera
- Sony Hi-MD minidisc recorder/player
- Sony small stereo microphone
- LCD video projector
- Regular-definition 21" color television
- Various audio/video cables
- Box of conversation cards from the board game 'Pharoah's Bluff'
- Couch (3-person/approx. 2m wide)
- 3 Filing cabinets 1.5m in height, 2 used as a sight barrier, one as a projector stand
- 2 Conference-type plastic chairs, used as speaker stands

- Large table used to hold the laptop, amplifier and television
- Small table used to hold the conversation cards
- Office chair for experimenter seating

Arrangement of Apparatus and Experimenter/Subjects

The subjects were seated side-by-side on the couch with enough room to move in contact or away from each other. In front of the couch, within easy reach of the subjects, was placed the table. The conversation cards were placed on the table. Immediately behind the couch, the projector stand was placed, with the projector resting on top of it. The projector settings were adjusted to project an image (approximately 2m wide by 1.5m high with bottom edge 2m above floor level) onto the blank white wall opposite the subjects (approximately 4m from the subjects). The microphone was attached to the vertical surface of the projector stand nearest to the subjects, at the subject's approximate head level. The speakers were placed on two chairs, located approximately 1.5m apart and 2.5m in front of the subjects, with each speaker oriented to point directly at the nearest subject. The video camera was mounted on the tripod and placed between the speakers, approximately 3m from the subjects.

The experimenter, system laptop, stereo amplifier and video monitor were located to one side of the couch, blocked from the subjects' view by a wall of 2 filing cabinets. The experimenter could hear the speech and system audio through the room and could watch the subjects via video camera feed to the video monitor.

Figure 11-2 below shows the arrangement of the experiment.

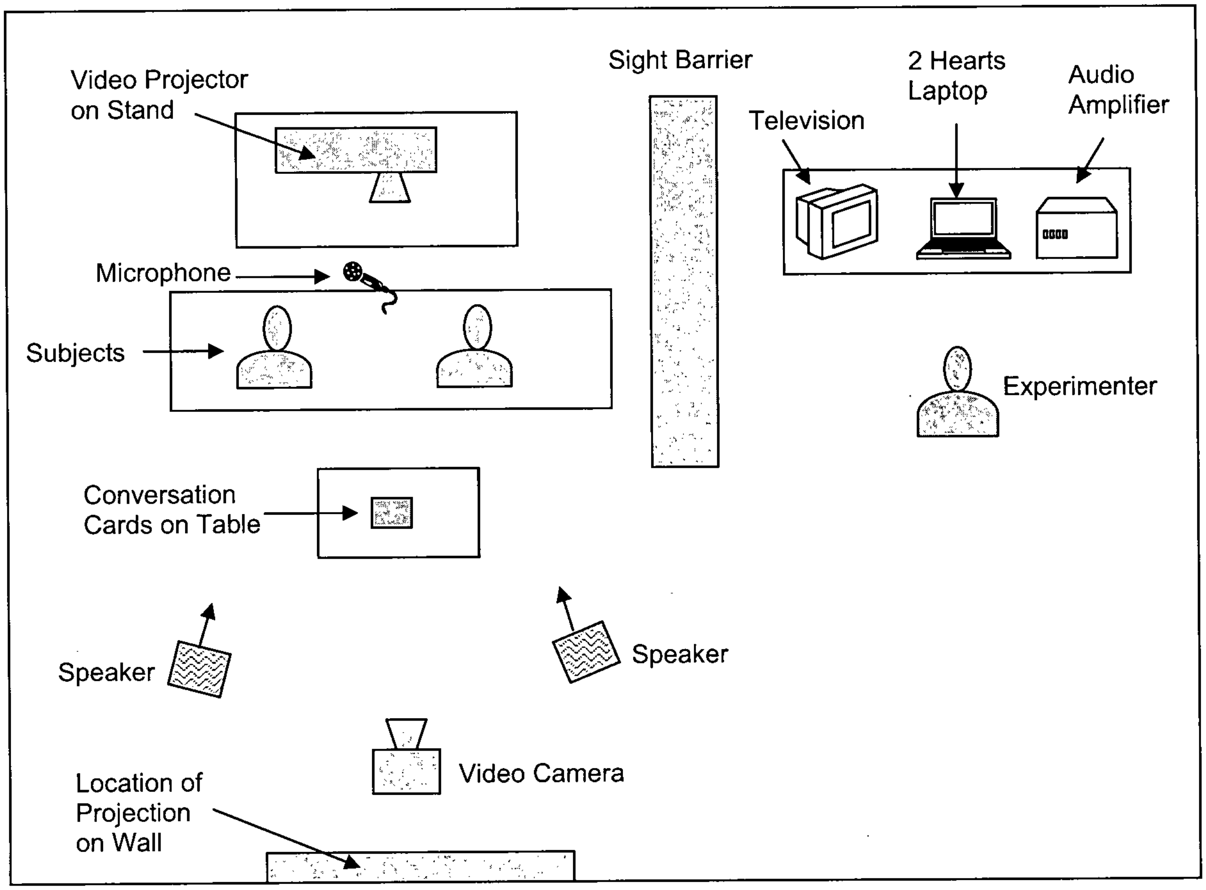


Figure 11-2. Arrangement Diagram of Apparatus and Experimenter/Subjects. (Not to scale)

Testing Procedures

The procedures for 2Hearts user testing are described in the order in which they were performed for each pair of subjects.

Preliminary

Subjects had read relevant information from the consent form on the experiment information web page prior to volunteering for the experiment (see Appendix D for consent-form/website text). Subjects read and signed the consent form in the experiment room, prior to the start of the experiment. The experimenter described how to wear the sensor and each subject was given a private room in which to put on the sensor. To ensure that subjects were comfortable with the social and system interaction during the experiment, subjects were instructed that they could stop at any moment. The experimenter remained in the room at all times to ensure that they could be contacted immediately. Amplification settings for the system in the testing room were fixed before the first user trial to create a volume level judged to be appropriate as clearly audible background music.

Resting Task

Subjects were told to sit comfortably, not touching each other, facing in parallel with each other. Subjects were told to close their eyes, be still and relax. They were told not to specifically meditate or do any relaxation exercises/techniques, but to just relax normally. Once subjects were relaxing, the period of heart data recording was begun. Resting task duration was 5 minutes.

Listening Task (including Questionnaire Part I)

Subjects were told that they were going to listen to pairs of sound samples and rate/describe the samples on a questionnaire. The subjects were shown the questionnaire sheet for Part I, and the technique for rating samples was explained.
They were instructed that each pair would consist of 2 one-minute audio segments, with a silent pause between the segments. Subjects were told to continue sitting comfortably in the same positions as the previous task, with eyes closed while listening.

Subjects were told that the first pair was to be played, and then the first pair of pre-recorded audio samples was played from minidisc, with a gap of approximately 10s between each sample in the pair. Subjects were instructed to open their eyes and given the questionnaire sheets to fill out regarding the first audio pair. The same listening/questionnaire procedure was followed for the 3 remaining audio-segment pairs.

Subject Introduction to Heart Feedback Audio

Beginning with one subject arbitrarily selected, each subject was introduced to the real-time audio feedback display of the 2Hearts system in the following way: the system audio was muted except for the varying-pitch tone sequence corresponding to the subject's heart data, and the subject was informed that each note was triggered by a heartbeat and that the pitch of the notes corresponded to the momentary heart rate. The subject was asked to observe the heart/audio connection using techniques of their choice such as feeling their wrist pulse, standing and moving, breathing fast or slow, relaxing, etc. This free-form exploration continued until the subject said that they understood the connection between heartbeat and audio.

Once both subjects were introduced to their own heartbeat displays as described, the two heartrate-controlled audio streams were turned on at the same time. Subjects were asked to observe how the two audio streams were differentiated by timbre quality and stereo separation, and asked to try and follow their own heart signal in the combined audio display. This period continued for roughly 2 minutes.

Conversation Task

Subjects were told that they were going to engage in a conversation task involving talking to the other person with the system audio feedback active. They were told that the system would possibly use different settings in each phase, and that in the different settings some of the audio information might be more or less related to the heartbeat sensor data. They were told that they would be discussing topics found on conversation cards.

Subjects were shown an example card and it was explained that each card contained 4 conversation questions, one of which to be chosen at will. Subjects were told that they were to alternate with each other in drawing the card and choosing the question. One subject was assigned by random coin-toss as the first card-reader. Subjects were told to discuss the selected topic for a length of time that seemed natural, up to a limit of roughly

2 minutes, after which they were to draw the next card. They were told that it was acceptable to comment on or discuss the system display and the topic of heart rates, but to generally focus on the given conversation topics. Subjects were told that they would be asked to provide information about the nature of their experience during the conversation on questionnaire forms.

Once subjects agreed that they understood the instructions, the system audio was engaged and the subjects were told to begin the first conversation phase. There were four 10-minute phases. After each phase, there was a pause/rest period of approximately 3 minutes, during which subjects were allowed to make notes regarding their experiences during that phase. Subjects were not shown the related questionnaire form until all phases of the conversation task were complete.

Questionnaire Parts II and III

After all phases of the conversation task were complete, subjects were given part II of the questionnaire to fill out, which involved rating and describing the musical and social experiences during each conversation phase. Subjects were allowed to consult the notes they made during the conversation task while filling out this part of the questionnaire.

After completing part II of the questionnaire, subjects were asked to fill out part III, which involved some subject information such as musical background, and general opinions/comments on the overall experience of using the 2Hearts system.

Control Game

Subjects were told that they were going to perform a game/challenge type task where the goal in each phase would be for one user (the 'controller') to influence the heart rate of the other (the 'target') in one specified direction. The subjects were informed that the controller was allowed to employ any desired technique of behaving or speaking to influence the target, except that the controller was not to physically contact the target or say things likely to offend the target. Users were told that the target user should not actively aid or resist the attempts at influence, but should rather try to react naturally to the speech/behaviour of the other person, and that the target should remain seated and relatively motionless.

Once subjects indicated that they understood the instructions, the task was started- the phases were performed consecutively, with no pauses between phases. Immediately at the beginning of each phase, the experimenter verbally indicated the phase goal to subjects using a phrase of the form, “Now, *raise* Alice’s heart rate” or “Now, *lower* Bob’s heart rate” (not actual subject names). The experimenter made an effort to say each phrase with a similar inflection and tone of voice: distinctly and quite loud, but with a neutral, calm affect.

After Testing

Subjects were debriefed as to the reasons for the study and allowed to ask questions about the study, which were answered by the experimenter. Subjects were offered the opportunity to be notified when the study was completed and to receive the results of the study.

Summary of Data Collection Methods

For each task, heartbeat sensor data for each task and subject was recorded onto hard disk using the 2Hearts software. For the subject-subject interaction tasks, video of the subjects along with synchronized audio from the 2Hearts system was recorded using the video camera, and room audio including subject speech was recorded on the minidisc recorder, using the stereo microphone. A regular-definition 21” color television was used to let the experimenter view the video camera signal in real-time, and make unstructured written observations regarding subject interaction and 2Hearts system performance. Written questionnaire data was also collected.

12. Analysis and Discussion of 2Hearts User Testing

This chapter presents observations, data analysis and discussion of results, grouped by Task. Statistical confidence intervals, t-tests and one-way analyses-of-variance were performed using the software package GraphPad Prism, while two-way analyses-of-variance were performed with Matlab version 7.0.4 (for Microsoft Windows) using the function "rm_anova2" as downloaded from the Matlab Central exchange website [96].

12.1 Observations

The following observations were made during 2Hearts user testing.

- In the training phase, subjects quickly understood the heart-audio connection and typically expressed enjoyment at hearing their biosignals in this way
- In the conversation task, it was realized that the physical motion of bending toward the table to pick up a conversation card could be influencing HR due to exertion
- In the conversation task, it seemed apparent that notable HR changes occurred between the moments of reading a card, considering the questions, picking and reading one question, and waiting for the response. It is speculated that orienting-response type HR patterns were observed.
- In the conversation task, there were at least two episodes where one subject became especially excited/amused and a clear increase in HR was heard by the experimenter in system audio
- When the volume of the filtered-noise audio stream grew high, the sound was quite distracting and annoying to the experimenter
- Often during the control-game task, the 'controller-user' stood up from the couch and faced the 'target-user'; at these times the controller could not view the visual heart feedback

- During the control-game, strategies employed by controller-users to influence the targets' HR included: use of humor, direct commands/instructions, speaking in a relaxed or excited vocal tone, looking into/away-from the target's eyes, verbally commenting on the feedback, describing imagery or scenarios, and discussing provocative topics. In one case, instructions to breathe slowly were given (which may be not viewed as a socially relevant interaction).
- Subjects were generally enthusiastic about the control-game and not at a loss about what to do

12.2 Segmentation of Raw Data and Derivation of Variables

The data recorder implemented in 2Hearts- Version 3 saved heartbeat data to an ASCII text-format file as a series of time-of-beat readings, measured in pd audio processing ticks since program initiation (64 audio samples per tick at a sampling rate of 44.1kHz). Separate files contained the saved marker times which indicated the start and end of various segments of the experiment. Files were saved for each subject and trial, named by subject ID-code and trial number.

Matlab was used for the following analysis tasks. Text files were imported as data, and the per-phase data was re-ordered according to the experiment task ordering scheme. For each subject/phase data segment, the following values of interest were calculated:

- A time series of beat indications, measured in seconds relative to the beginning of the segment
- A global per-segment mean heart rate value, measured in BPM (equal to the total number of beats in the segment divided by segment duration in minutes)

- An evenly-resampled 100Hz instantaneous heart rate signal in BPM, having the following properties:
    - Any time point falling between two measured beat-times is assigned the rate value equal to the inverse of the time difference between those two beats
    - Any time point equal to a measured beat-time is assigned the rate value equal to the inverse of the time difference between that beat and the previous beat

(Expressed symbolically, this interpolation algorithm gives:

$$\text{If} \begin{cases} \mathbf{tbeat}_n < \mathbf{t} < \mathbf{tbeat}_{n+1} \quad \text{then} \quad \mathbf{IHR(t)} = \dfrac{60}{\mathbf{tbeat}_{n+1} - \mathbf{tbeat}_n} \\ \\ \mathbf{t} = \mathbf{tbeat}_n \quad \text{then} \quad \mathbf{IHR(t)} = \dfrac{60}{\mathbf{tbeat}_n - \mathbf{tbeat}_{n-1}} \end{cases}$$

where $\mathbf{t}$ = time in seconds , $\mathbf{tbeat}_n$ = time-of-beat reading with index n, in seconds and **IHR**(t) = instantaneous heart rate value in beats per minute.)

Each data set was visually inspected to ensure that the derivation calculations had given reasonable values, and it was observed that false sensor readings as indicated by outlying high or low heart rates were rare.

Figure 12-1 below shows an example of the interpolated heart rate data of a user during the conversation task with the system enabled (display scheme 1).

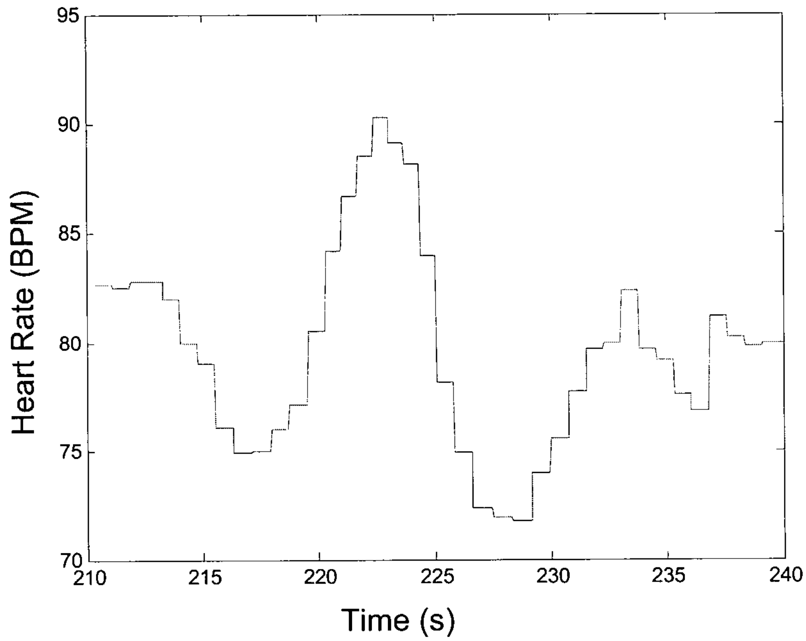


Figure 12-1. Example of Interpolated Heart Rate Graph. Heart Rate (in Beats Per Minute) vs. Time (in seconds). For one subject, during Conversation task, using audio display scheme Type 1.

For user responses entered on the continuous-scale questionnaire sections, the location of the users' hand-drawn slash marks on the 15cm-wide scale were converted to numerical data by human measurement. A custom measurement tool was used, 15cm-wide with 100 equal gradations, and marked with values 0.0 through 10.0. Thus, numerical values from 0 to 10 were obtained with 1 significant decimal place.

12.3 Task I Data Analysis (Listening Tasks)

Mean Heart Rate for Resting vs. Listening

It has been suggested previously that listening to music tends to raise the listener's heart rate, compared to sitting in silence, although the effect may depend on musical style (see Section 6.3). It was explored whether the 2Hearts system audio would cause such an effect, by comparing each subject's mean heart rate for the silent resting task with their mean heart rate for the combination of all the music-listening tasks.

Cumulative mean HR for the excerpt-listening tasks together was determined by taking for each subject the mean of the per-excerpt HR means. Mean HR from a 5-minute resting task was compared with mean HR from listening tasks of 8 minutes total. Figure 12-2 shows this comparison, illustrating within-subject change.

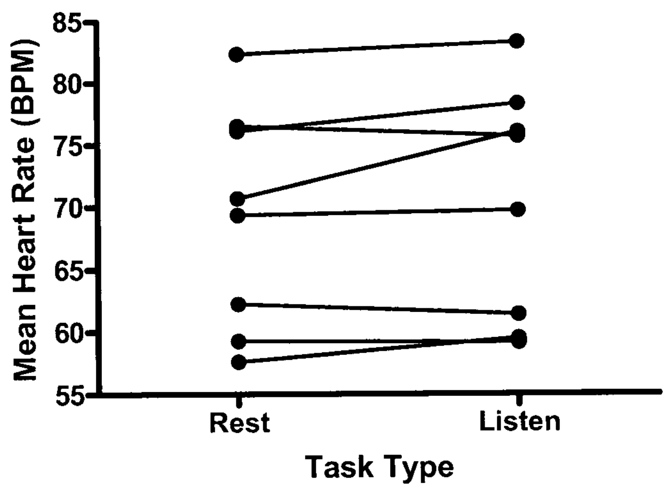


Figure 12-2. Mean Heart Rate for Resting vs. Listening Tasks. Eight subjects were used. A paired t-test failed to find a significant change in mean HR between the two conditions (p = 0.17).

The mean of differences was 1.1 BPM in the predicted direction (greater heart rate for the listening tasks). A paired t-test (two-tailed) was used to check for a significant within-subject change in mean HR between the two conditions. The test failed to show a significant change ($p = 0.17$).

Mean Heart Rate Compared Within Matched Pairs of Audio Excerpts

For each of the audio pairs, the within-subject change in mean HR between the two excerpts was analyzed. Figure 12-3 shows the data.

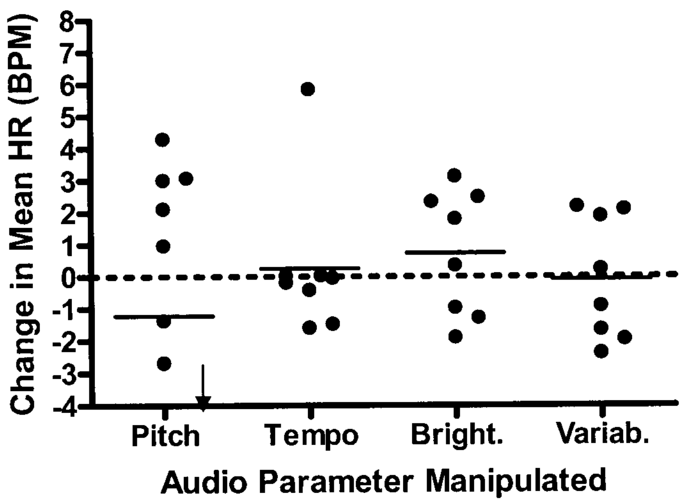


Figure 12-3. Change in Mean Heart Rate vs. Audio Parameter Manipulated. Eight subjects were used. Paired t-tests failed to show significant within-subject change between the two excerpts within each pair, for each of the 4 parameters pitch, tempo, brightness, and frequency variability The arrow points towards one outlier at -19.1 BPM.

For each parameter, a separate paired t-test (two-tailed) was conducted to check for a significant within-subject change in mean HR between the first and second excerpts. None of the excerpts lead to statistically significant changes in mean heart rate.

Mean Heart Rate Compared Between Pairs of Audio Excerpts

While analyzing the data, the question arose whether any difference in heart rate occurred between the different timbres/styles of audio excerpt (ie. between the excerpt pairs.) Figure 12-4 shows the response of each subject's HR to each grouped pair of audio excerpts.

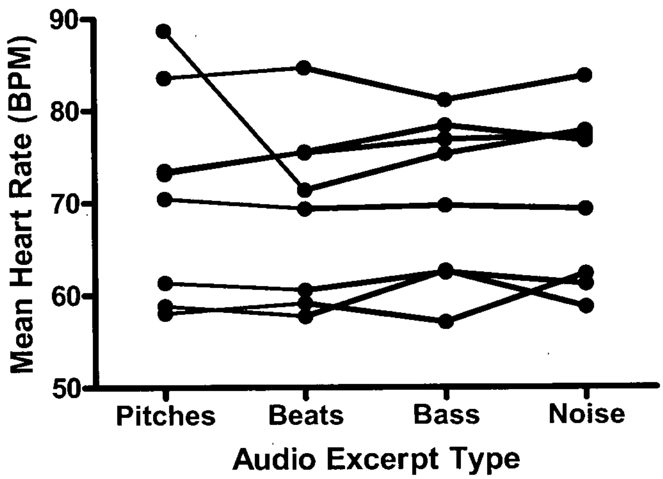


Figure 12-4. Mean Heart Rate for Four Types of Audio. Eight subjects were used. A one-way within-subjects ANOVA failed to show any significant change in mean HR related to the variation of the audio-type condition ($p = 0.68$).

A one-way within-subjects ANOVA was performed to check for a significant relationship between audio-type and HR. No significant relationship was found ($p = 0.68$).

Questionnaire Part I- Ratings Comparing Perceived 'Happiness' Within Excerpt Pairs

The subject ratings of 'energetic' difference within the 4 excerpt pairs are shown in Figure 12-5.

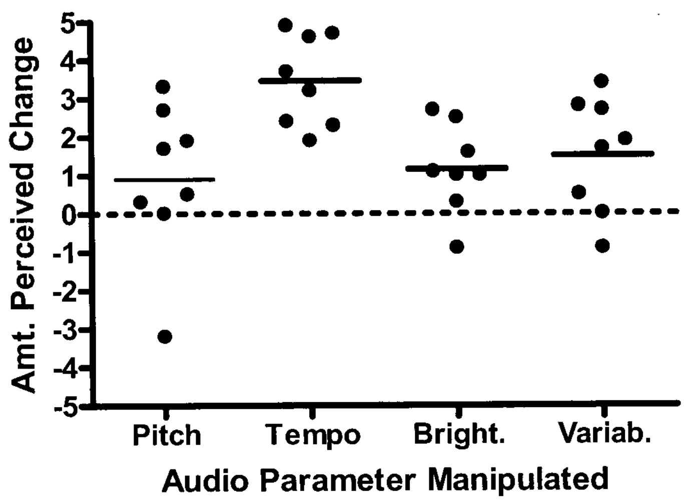


Figure 12-5. Change in Perception of 'Energy' for Four Manipulated Audio Parameters. Eight subjects were used. On the vertical scale, +5 represents that the excerpt with higher parameter-value was rated as extremely more energetic than the excerpt with low parameter-value; -5 represents that the excerpts contrasted extremely in the other direction.

To examine the significance of subjects' perception of change in 'energy' for different types of audio manipulation, 95% confidence intervals for the underlying mean change were calculated. Figure 12-6 shows the results.

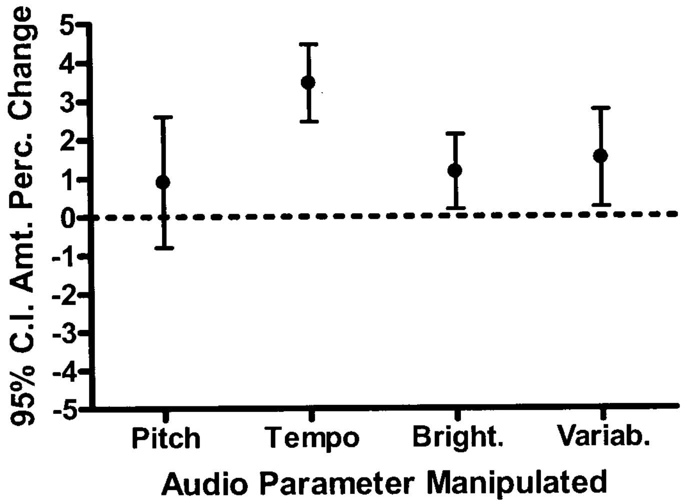


Figure 12-6. Confidence Intervals for the Change in Perception of 'Energy' for Four Manipulated Audio Parameters. Eight subjects were used. On the vertical scale, +5 represents that the excerpt with higher parameter-value was rated as extremely more energetic than the excerpt with low parameter-value; -5 represents that the excerpts contrasted extremely in the other direction. For tempo, brightness, and frequency-variability, the zero value indicating no change lies outside of the confidence interval.

The confidence intervals for pitch, tempo, brightness, and frequency-variability were calculated as (-0.80 to 2.6), (2.5 to 4.6), (0.19 to 2.1) and (0.25 to 2.8), respectively. Note that the CI's for tempo, brightness, and frequency-variability do not include the zero-value indicating no perceptual change.

Questionnaire Part I- Ratings Comparing Perceived 'Happiness' Within Excerpt Pairs

The subject ratings of 'happiness' difference within the 4 excerpt pairs are shown in Figure 12-7.

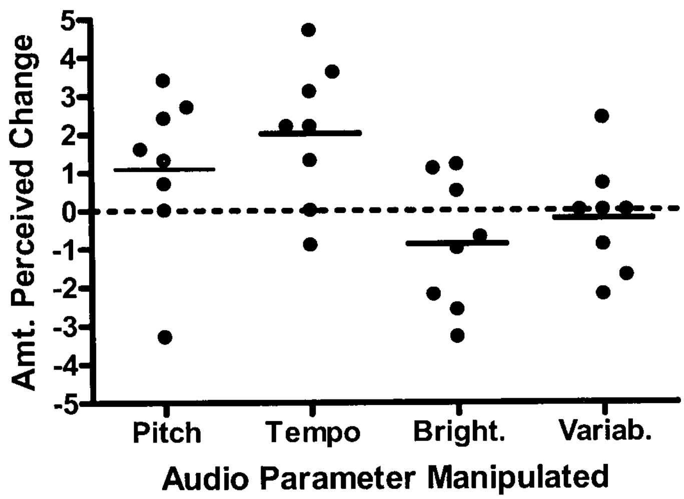


Figure 12-7. Change in Perception of 'Happiness' for Four Manipulated Audio Parameters. Eight subjects were used. On the vertical scale, +5 represents that the excerpt with higher parameter-value was rated as extremely more energetic than the excerpt with low parameter-value; -5 represents that the excerpts contrasted extremely in the other direction.

To examine the significance of subjects' perception of change in 'happiness' for different types of audio manipulation, 95% confidence intervals for the underlying mean change were calculated. Figure 12-8 shows the results.

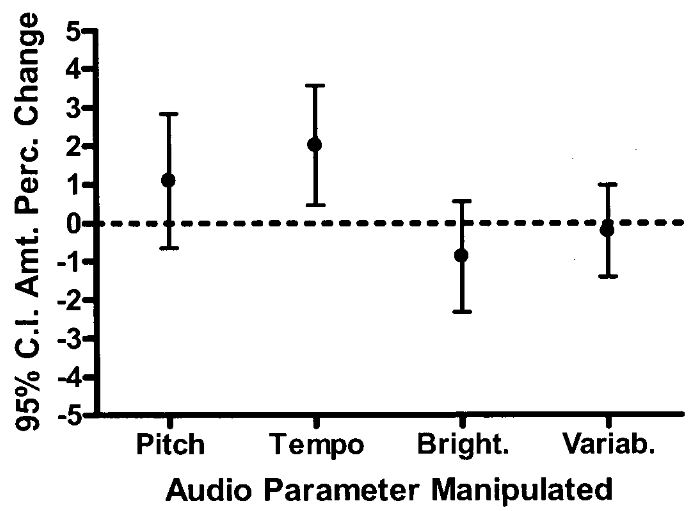


Figure 12-8. Confidence Intervals for the Change in Perception of 'Happiness' for Four Manipulated Audio Parameters. Eight subjects were used. On the vertical scale, +5 represents that the excerpt with higher parameter-value was rated as extremely more energetic than the excerpt with low parameter-value; -5 represents that the excerpts contrasted extremely in the other direction. For tempo, brightness, and frequency-variability, the zero value indicating no change lies outside of the confidence interval.

The confidence intervals for pitch, tempo, brightness, and frequency-variability were calculated as (-0.65 to 2.9), (0.48 to 3.6), (-2.3 to 0.56) and (-1.4 to 0.99), respectively. Note that only the CI's for tempo does not include the zero-value of no perceptual change.

Questionnaire Part I—Free Form Section Comparing Audio Excerpts

It was observed that the designed categories of 'How did each excerpt sound' and 'How did each excerpt make you feel' did not lead to 2 clearly distinguished response types- subjects gave both emotional and musical descriptions in each category (an interesting fact in itself.) Also, the format of describing each audio pair-member separately made it difficult to separate phrases meant to contrast the pair-members from phrases meant to apply to both members. On the other hand, many of the responses were colorful and evocative, and there seemed to be apparent differences between the descriptions used for different types of audio stream.

Therefore, it appeared that the data could be used informally to gain some insight into the different ways subjects experienced/conceptualized the musical connotations of the audio streams. For each audio-pair, all response categories were first combined; then, by qualitative inspection, some common types of response were identified.

The most common response types are listed for each excerpt type, along with the number of different subjects who made responses of that type, and example quotes. (These numbers are intended only as suggestive guidelines and not intended to be interpreted as statistically significant findings.)

'Pitch Sequence' Excerpts

- **Anxious/Uneasy (5 subjects):** "Cautious", "It sounds eerie", "A little creepy".
- **Sad/Depressed (3 subjects):** "was a real downer", "made me feel lower, 'depressed' ", "it sounds mournful"

'Percussive Beats' Excerpts

- **Aroused with Positive Affect (6 subjects):** "Happy- like dancing", "made me feel a bit more heightened [...] I like it", "felt energy, felt a bit like dancing".
- **Relaxed with Positive Affect (5 subjects):** "I felt kind of relaxed and in the groove", "soothed and sleepy", "balanced, relaxed".

'Bassline' Excerpts

- **Irritated/Disturbed (4 subjects):** "Confused and somewhat irritated", "it's high and irritating", "The buzzing was a little disconcerting".
- **Comical (4 subjects):** "I began to find it comical and bumbling", "the sound was a bit humorous", "silly, light hearted".

'Noise' Excerpts

- **Like Wind (5 subjects):** "the sound is a desert sandstorm", "like a harsh wind blowing", "wind was fiercely blowing me around".
- **Anxious/Uneasy(4 subjects):** "Anxious, uneasy", "stressed out, anxious, guarded", "anxious at the very end"

Thus the responses appeared to indicate some distinct categories of description related to the timbre/behaviour of the different audio streams.

An aspect of the responses that has not been quantified is that when comparing the within-pair changes for 'pitch sequence' and 'percussive beats', subjects seemed to employ similar terms (such as 'higher'/'higher pitch' for pitch change and 'faster'/'sped up' for tempo change) whereas changes in brightness and frequency variability were described using a heterogeneous collection of terms (including "more resonant", "had a bit more… distort? Fuzz?" for brightness and "more swirly", "faster and more like a spring breeze" for frequency variability).

Another notable aspect of the responses was that many responses were framed in terms of the subject's own emotions (e.g. "I felt kind of relaxed") while many other responses were framed in terms of describing an external object (e.g. "I began to find it comical and bumbling. It's cute.")

A final observation is that in several cases, unpredictable or jumpy audio behaviour was associated with negative valence connotations.

12.4 Task I Discussion

Mean HR for Resting vs. Listening

Regarding the comparison of resting vs. listening heart rates, the lack of a significant finding gives some idea that a bias toward arousal when listening to music may not be an overwhelming factor in the 2Hearts system design. This fact would be helpful if true, since the SBFS needs to effectively communicate both low- and high-arousal information. Task I was not directly comparable to any particular previous study which found a general increase in heart rate caused by music listening. Therefore the current finding is not a failure to replicate; rather, it could be that the 2Hearts audio streams manage to be neutral rather than exciting or calming. To demonstrate this would require the addition of control audio-streams which cause measurably increased/decreased heart rate.

Consideration of these results has led to an increased appreciation of the advice of Harrer and Harrer, who caution that experimental setting, task ordering, musical style and subject attitude may have a large effect in this stype of study [80]. In particular, it is now realized that subjects might have been aroused during the 'resting' task due to the unfamiliar experiment room and the fact that they had not yet experienced any audio or feedback and may have felt anticipatory/nervous as a result.

Mean HR for Audio Excerpts

Having not found a significant difference between no-music and music, it appeared doubtful that the briefer and more subtle parameter-change tests would show significant differences. Indeed, no significant differences were found within or between excerpt pairs. This could indicate a lack of statistical power in the test-- given the wide spectrum of thought/mood fluctuation people may experience in a given minute, regardless of any musical stimulation, it is possible that the effect of musical variation is not distinguishable with the sample size used. Also, the fact that subjects were instructed to listen consciously/analytically during the listening task may have prevented emotional involvement in the music and prevented the music from maximally influencing their physiology. (In other words the task instructions might have strengthened S→C at the expense of S→U.)

On the other hand, the lack of strong results could indicate that musical variation within the normal range produced by 2Hearts-Version 3 is not suited for producing dramatic physiological reactions within the minute-by-minute time-scale examined.. It is suggested that future listening tests of 2Hearts audio should incorporate a step where previous results are replicated, showing a significant variation of heart behaviour with musical condition. This would ensure that the general test conditions and statistical power of the experiment are adequate to find music-induced heart responses if such responses are present. The study [82] has been identified as a useful candidate for future replication. If 2Hearts audio parameters are in fact less effective than other schemes tried in the past, past schemes could give direction as to improving the parameter mapping

Questionnaire Part I- Ratings

Table 10-1 can be consulted to review the designed directional bias of the 2Hearts display mappings. For the audio parameters of tempo, brightness, and frequency-variability, subjects' ratings of 'less/more energetic' increased significantly with parameter increase. Similarly, the ratings of 'happier' increased significantly with tempo, and did not show significant change for other parameters. Table 12-1 below relates the significant results to desired bias design goals of the 2Hearts system:

Stream Description	Parameter Tested	Desired Relative Bias	Measured Relative Bias
'Lead' Pitch Sequences	Pitch	Higher pitch = higher arousal	None significant
'Bassline'	Brightness	Brighter = higher arousal, brighter = negative valence	Brighter = higher arousal
'Noise'	Centre Frequency Variability (not an explicit design parameter)	Not known in advance	More variable = higher arousal
'Drumbeat'	Looping Tempo	Faster beat = higher arousal	Faster beat = higher arousal, faster beat = positive valence

Table 12-1. Summary of Self-Report Test Showing Users' Perception of Display Parameters.

Overall, these results suggest that many of the parameter selection choices need to be re-thought or manipulated through timbre design in order to carry the designed directional bias. It is postulated that these measurements of the directional bias DCC correspond more directly with the S→C pathway than the S→U pathway, as they were mediated through conscious self-report.

Questionnaire Part I- Free Form Section

As mentioned, the format of the questionnaire led to difficulty in evaluating directional bias information from the user reports; however the observation of distinct user description-patterns for the different audio streams provides findings regarding the DCC of absolute bias. In particular, the pitch-sequence audio stream was found to convey anxiety, and the bassline audio stream was perceived as comical or irritating. Thus these streams conveyed unwanted absolute bias. Such findings give useful direction for improvement of the 2Hearts audio mapping and timbre design.

Possible sources of perceived 'anxiety' in the pitch-sequence stream were the dissonance resulting from random non-scaled semitone relationships, and the unpredictable behaviour of the random pitch increase/decreases. (This latter source of anxiety-perception may be ameliorated if real HR input data is used and the data varies more regularly than the 2Hearts input-simulation algorithm.)

The format of this part of the questionnaire could be improved to better separate relative or 'comparing' words with absolute or 'describing' words, and to better separate perceptions of objective sound variables from perceptions of emotional impact. Perhaps a standardized system of ranking/classifying different descriptor words could be found.

Task I General Discussion

This task was constructed with several arbitrary/estimated choices of constants, especially the pairs of contrasting minimum/maximum values set for audio-parameter variation. The relevancy of the task could be improved in the future by using real pre-recorded HR data to drive the audio display, and by varying between-condition parameters within ranges measured from the actual HR data.

12.5 Task II Data Analysis (Conversation Task)

Task II Heart Rates, Between Condition Within Subject

Data from the conversation task were analyzed to determine if the system audio condition had any influence on subjects' mean HR. Figure 12-9 shows grouped HR data from the 8 subjects for each audio condition.

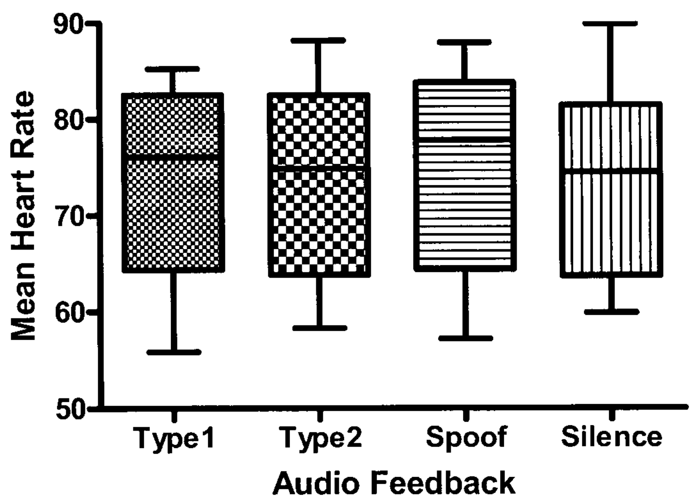


Figure 12-9. Mean Heart Rate for Four Audio Feedback Conditions. Eight subjects were used. This 'box and whiskers' plot indicates the mean (horizontal line), quartiles (box) and extreme values (whiskers). A one-way within-subjects ANOVA found no significant relationship between mean HR and audio-feedback type (p = 0.54).

Task II Heart Correlations Between Conditions for Subject-Pairs

To test for the presence of synchronized HR behaviour between users, Matlab was used to calculate the Pearson's correlation between paired users' HR signals, for each segment of the conversation task. Figure 12-10 shows the resulting correlation values for each audio-feedback condition.

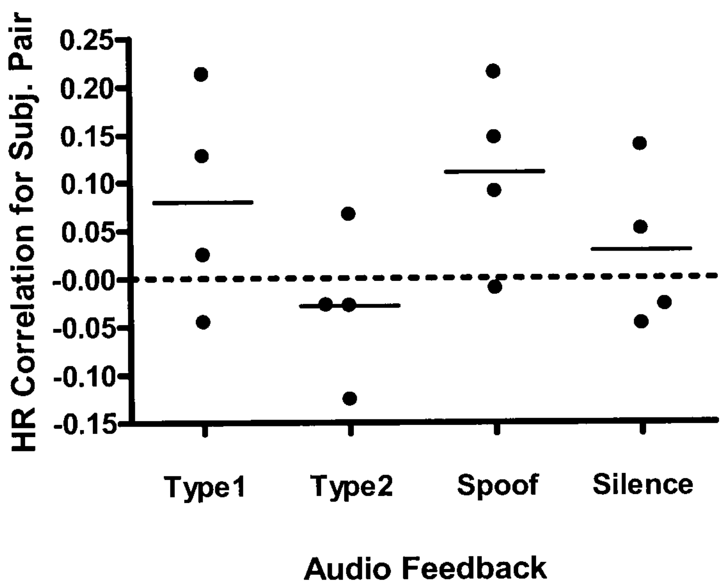


Figure 12-10. User-User Heart Rate Correlation for Four Audio Feedback Conditions. Correlations were calculated based on 4 pairs of users. A one-way ANOVA found no significant relationship between audio condition and HR correlation ($p = 0.22$).

Preliminary plotting of the data suggested no within-subject (or rather, within subject-pair) trends in the data, so a regular one-way ANOVA was performed to test for significant relationship between audio condition and correlation. No significant relationship was found ($p = 0.22$).

Due to the relatively low number of data points (resulting from the necessity of combining data in pairs) it seems likely that this test lacked statistical power. In particular, higher mean correlation is noted for the audio-type 1 and audio-spoof conditions. It is noted that a common factor between the audio-type 1 and audio-spoof conditions is that these conditions both employed the 'bassline' and 'noise' audio streams while the other conditions did not. This is discussed further in Section 12.3.

In order to explore user-user correlation as a generally useful measure of HR behaviour, the correlation measurements for each subject-pair and task-segment were considered as a group. Figure 12-11 illustrates this combined distribution.

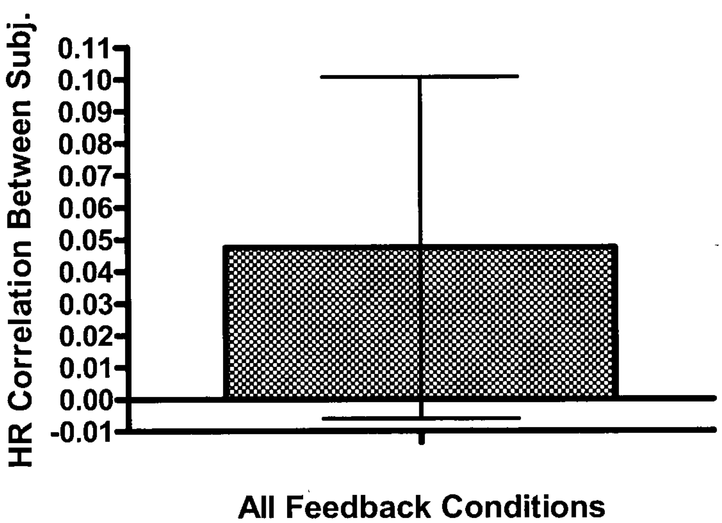


Figure 12-11. Distribution of Results for All Heart Rate Correlation Measurements. Four subject-pairs x 4 conditions were considered. The bar indicates mean correlation and the line-segment indicates 95% confidence interval of the underlying mean.

A 95% confidence interval was calculated for the underlying mean of the combined correlation measurements, giving (-0.01 to 0.10). The value of zero correlation is not excluded from this interval, indicating it cannot be discounted that this type of audio/social context may have no tendency to produce HR-HR correlation.

Questionnaire Ratings Comparing the Audio Conditions

Subjects' ratings comparing the 4 audio-conditions in the conversation task were examined. Preliminary examination of the data did not suggest any significant trends. For each of the 6 questionnaire categories, a within-subject one-way ANOVA was performed to compare subject's responses to the 3 audio conditions of audio-type 1, audio-type 2, and spoof (the silent condition was excluded since questions referred to 'the system audio'). No significant relationships were found.

Freeform Questionnaire Findings Regarding Conversation Task

Inspection of the free-form responses describing the conversation task phases showed a wide variety of responses, with many responses describing specific social events particular to a given subject-pair and conversation. A grouping or rating scheme for these responses has not yet been identified.

12.6 Task II Findings and Discussion

HR and HR-HR Correlation

None of the HR-based measures varied between audio-feedback conditions in a significant way. Thus it could not be shown that the current 2Hearts-Version 3 design provides any significant augmentation along pathway S→U→V when used in a 'background music' context.

One trend was observed which seemed to nearly meet significance: an elevation in HR-HR correlation for audio conditions 'Type 1' and 'Spoof'. The common factor between these conditions is the presence of the 'bassline' and 'noise' correlation displays; it is speculated that when these audio streams became loud/bright and noticeable, both subject's HRs were influenced to increase.

Considering causes of HR-HR correlation in this way highlights the fact that for SBFS's, it is difficult to distinguish the effects of a 'true' socially-created HR synchronizing event from an event caused by both users perceiving the same audio stimulus. In other words, even a faulty or random-behaving SBFS can increase correlation merely by creating arousing or relaxing sonic events, regardless of input data. This problem of distinction requires further study.

The total task-wide measure of HR-HR correlation had a mean of greater than zero, and came close to indicating an underlying population mean of greater than zero. Further study of this measure in more/longer tasks could be used to support the idea of HR-HR correlation as a measurable phenomenon.

Questionnaire Part II

The self-report ratings failed to show any significant relationship with audio-feedback conditions, in any of the measured categories. This indicates that subjects could not distinguish real feedback and spoof conditions, and therefore the total system performance with respect to pathway S→C was not verified. Perhaps the construction of the spoof condition was inadequate to decorrelate input/audio-feedback; this should be examined further.

The free-form section results highlighted the wide variation in different social situations and moods experienced by the users; this points out the fact that even a well-designed SBFS may require performance tests of high statistical power in order to overcome the random social variation of affective variables.

General Discussion

In general, this task did not lead to strong indications of system performance. Two potential reasons for this are problems with system construction, and problems with task design/suitability.

Task I results give many indications of problems in bias of the 2Hearts display channels; these problems could have led to poor perception of system feedback in general. Unwanted audio connotations of irritation/anxiety may have overwhelmed the correct conscious/ unconscious interpretation of patterns in the audio feedback. It is also speculated that the information in the pitch-sequences indicating per-user HR competed directly for attentional bandwidth with users' focus on conversational speech sounds- this competition would be exacerbated by the similar hi-mid frequency ranges of system sounds and speech sounds.

Regarding task suitability, users were told to keep returning their attention to the conversation card topics; in hindsight this instruction may have caused a disadvantage for the perception of system audio. The general strategy of testing the 2Hearts system in a heavily topic/conversation oriented task may be misguided; system design was undertaken mostly with the 'co-navigate' paradigm in mind which involves a greater focus on free-form playing and listening/commenting on system audio. Additionally, the task was not designed to account for longer-term learning requirements which may be necessary to optimally use an SBFS system: perhaps future tasks should employ longer training periods over multiple sessions.

Regarding the experimental design of this task, it remains promising but needs improvement. Observationally, the task did succeed in creating a standardized social scenario where subjects were able to quickly begin interacting and keep interacting in a similar way for 40 minutes; such a task may still be of use in testing certain aspects of SBFS design. A great improvement in the task would be achieved by finding independently measurable social markers/indicators linked to empathy or other affect-related variables; such secondary independent measures would make it easier to quantify SBFS success at providing socially relevant feedback. Another aspect of the task needing improvement is that subjects in many cases found it difficult to remember in detail their exact experiences; perhaps methods could be employed where subjects viewed and rated the videotaped interaction at a later time.

12.7 Task III Data Analysis (Control-Game Task)

Relationship of Users' Heart Rates to HR Goal and Feedback Condition

For the control-game task, HR data was separated into 2 sets: data from users when acting as 'controller', and data from users when acting as 'target'. The data were then examined with respect to the 2 independent variables of audio condition and HR goal.

Figures 12-2 and 12-3 below show the distribution of HR results for 'controller' users and 'target' users, respectively.

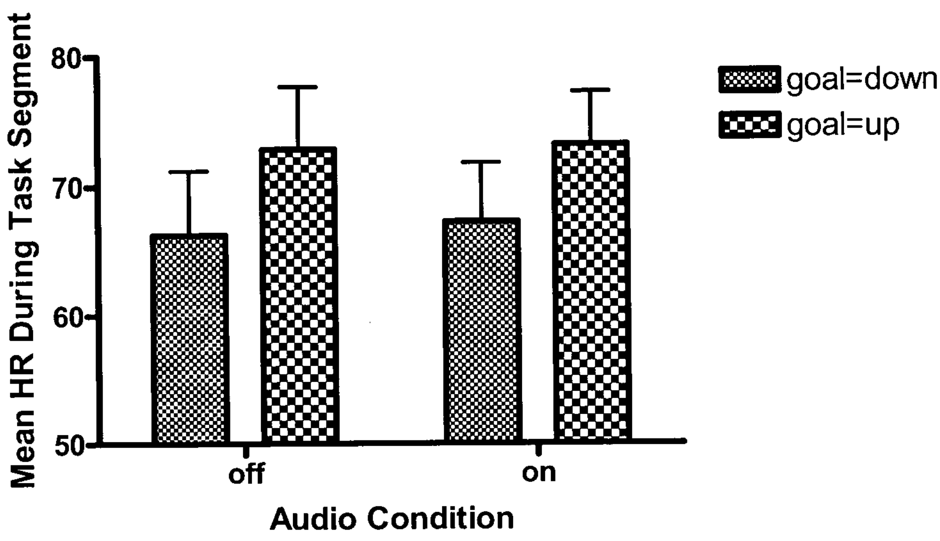


Figure 12-12. Mean Heart Rate of Subjects Acting as 'Target'. Shown for values of the conditions HR Goal = [up, down] and Audio Feedback = [off, on]. Six subjects were used. Bars indicate mean HR and error-bars indicate standard error.

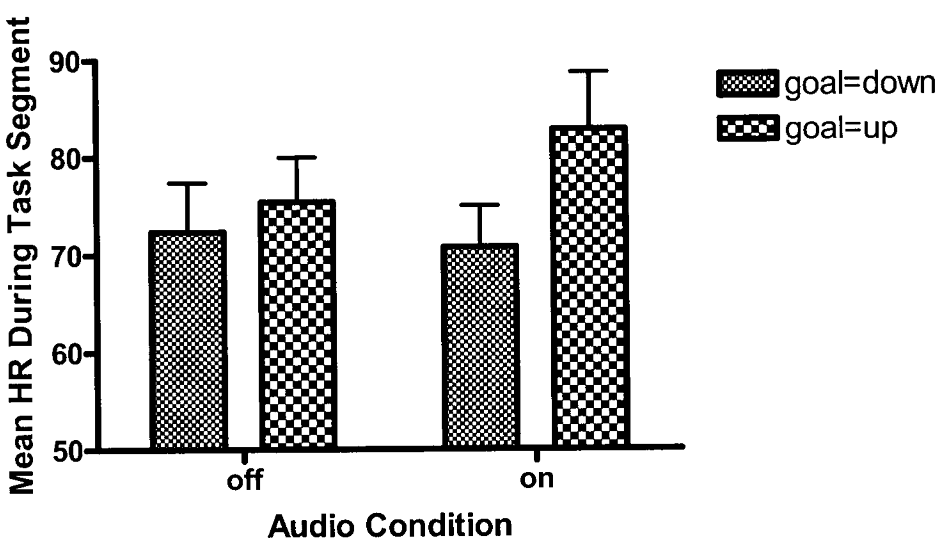


Figure 12-13. Mean Heart Rate of Subjects Acting as 'Controller'. Shown for values of the conditions HR Goal = [up, down] and Audio Feedback = [off, on]. Six subjects were used. Bars indicate mean HR and error-bars indicate standard error.

For each of the two data sets, a two-way within-subjects ANOVA was performed. In each case, a significant interaction was found between HR goal and mean HR. No significant interaction was found between audio condition and mean HR, or between the two condition variables; however some of these relationships were nearly significant. Table 12-2 below summarizes these results.

Role of Subject	**HR Goal / Mean HR Relationship**	**Audio State / Mean HR Relationship**	**HR Goal / Audio State Relationship**
Target	P = 0.023	P = 0.64	P = 0.72
Controller	P = 0.011	P = 0.066	P = 0.073

Table 12-2. Summarized Two-Way ANOVA Results for the Effect of HR Goal and Feedback Condition on Mean Heart Rate.

12.8 Task III Findings and Discussion

For users acting as 'targets of control', it was shown that HR was significantly higher when the segment-goal was 'raise HR'. This does not relate directly to SBFS performance, but does confirm an important working assumption for general SBFS applications: the assumption that one person can act intentionally in a social way to control the HR of another. This result indicates that a heart-sensor-based musical instrument could be successfully controlled to some degree through social interaction techniques.

A similar finding was made for users acting as 'controller'; however, this could have simply resulted from the physical activity taken by the controller-user (whereas the target-user remained motionless.) An interesting relationship for controller-users, that appeared nearly significant, was a positive correlation between presence-of-audio-feedback and agreement of controller-HR with goal-HR-direction. It can be speculated that since the controller-user often had their back turned to the visual feedback, they became excited by the increasing audio feedback during 'raise HR' segments.

For this task design, the lack of significant measured correlation between the two independent variables means it was not shown that the system provided augmentation of the 'control' SBFS paradigm. In hindsight, it was too ambitious to attempt to show superiority of system audio feedback over a video-feedback control condition. A better choice would have been to first establish the effects of feedback in general for this task.

The control-game shows promise as a SBFS testing technique, since users were able to successfully perform it in the expected manner. One flaw of the task is that both users are likely aware of the hypothesis and maybe subconsciously attempt to support it. Another limitation of the task as designed is that it would not distinguish between two possible causes for an enhancement in the measured/goal HR agreement—the cause of HR feedback from target-user to target-user and the cause of HR feedback (and thus control enhancement) from target-user to controller-user. In signal-flow terms, (given user 1 as controller and user 2 as target) the pathways U2→S→U2 and U2→S→U1/C1 are not distinguished by this task. However, for future experiments, it could be possible to deliver audio in isolation to one user at a time in order to make such a distinction.

12.9 User-Experience Questionnaire Analysis

Rated Attitudes toward General Experience Using 2Hearts System

Subjects' responses to the 6 five-point-scale questions in Questionnaire Part III were examined. Summary statistics for the responses were calculated. Table 12-3 below summarizes the user responses for each question.

Questionnaire Statement	Mean Response (Range 0-5)	Std. Dev. of Response
Liked System Sounds	3.8	1.3
Music was distracting	3.5	1.1
Sensor was distracting	2.3	1.2
Enjoyed System Overall	4.4	0.7
Got to know other subj.	4.1	0.4
System helped get to know	3.3	0.7

Table 12-3. Summary of Subjects' Ratings of Six Statements Regarding 2Hearts System. Eight subjects were used. The scale used was a 5-point scale with 0=strongly disagree, 5=strongly agree. See Appendix E for full wording of statements.

Figure 12-14 illustrates the distribution of responses.

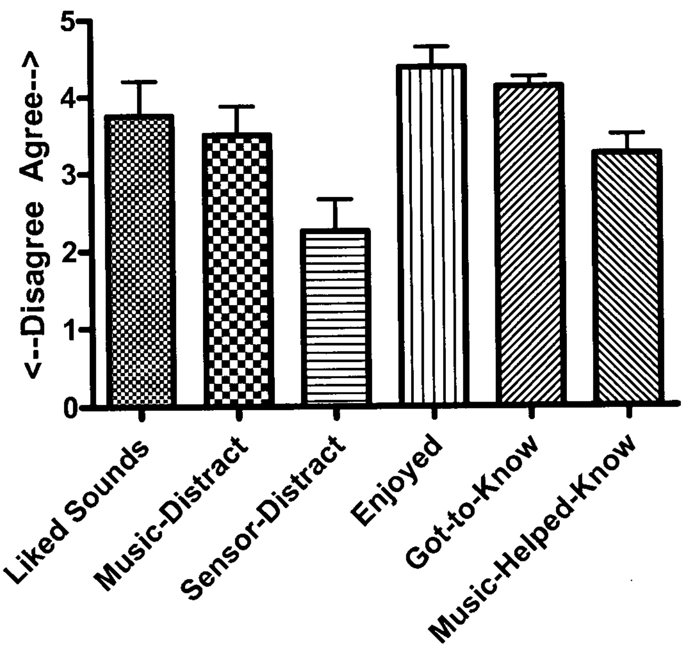


Figure 12-14. Subjects' Ratings of Six Statements Regarding 2Hearts System Use. Eight subjects were used. Bar height indicates mean and error-bar height indicates standard error. (For full text of statements see Appendix E.)

Freeform Questionnaire Responses Regarding General Experience Using 2Hearts System

Subject responses in this section were grouped and analyzed informally. Major response types and individually notable responses are presented here for the 3 questions.

1. Question: 'please describe what you liked/didn't like about the sounds produced by the system'

There were many interesting responses for this question. Responses within 3 categories relevant to SBFS design are listed:

Specific Opinions Regarding Audio Design

- "I liked the beat and the fuzzy/distorted sound"
- "The distorted feel to some of the sounds can make my skin crawl"
- "Didn't like the heavy dark sounds"
- "I liked their ambiance"
- "I wished they were more textured, rhythmic. This would make it more "musical" to my ears."
- "The higher sound was distracting and annoying. It sounded too unpredictable with the way the pitch was changing and it sounded like it was coming in and out of focus"
- "I liked the second set of sounds because I felt like there was a solid beat I could relate to."
- "Liked the deeper "bassier" sounds"
- "[Liked the] variety of samples"
- "Overall sound quality was excellent [...] interesting and intriguing soundscapes"

Distraction/Interference with Conversation

- "The heavy dark sounds [...] interfered with conversation"
- "[The sounds] sometimes were distracting"
- "I can't decide if it was too unobtrusive"

Feeling/Understanding of Connection with Sounds

- "The second set of sounds [...] seemed to support the rapport between me and the other [subject]."
- "I liked it when I could hear my heartbeat in a pleasant sound and I disliked hearing unpleasant sounds in connection with my heartbeat."
- "I also can't decide if the beat ←→ heartbeat connection was too obvious"
- "Didn't like not knowing what/why the sounds reacted the way they did"
- "Would have liked to feel capable of controlling sound parameters more somehow(??)"

2. Question: 'Would you choose to use this type of system again? What did you like/dislike about using the system'

The results for this question were fewer, and fell into categories less obviously. A few notable responses were:

- "Maybe once or twice for therapy to increase my awareness of what raises my heart rate"
- "I'd like it if users could select what type of sound will be generated by or triggered by their own heart"
- "I liked taking my pulse while hearing the system react to it"
- "A unique experience"

3. Question: 'In what context (place, time, setting) do you think it would be most helpful or enjoyable to use this type of system?'

The results for this question fell naturally into several groups, along with some miscellaneous responses. The following list summarizes repeated responses and miscellaneous responses:

- At a party (4 subjects)
- At home with friends (3 subjects)
- Performance/multimedia art (3 subjects)
- Spiritual use/meditation (2 subjects)
- Counselling/therapy (2 subjects)
- Miscellaneous answers:
  "Two people in different locations"

"It would be excellent to use during sexual encounters"
"At a mental spa"
"Educational or medical purposes"
"May work well at office meeting for brainstorming sessions"
"A good ice-breaker for people who need to work or interact with others fairly closely in a short period of time"

12.10 User-Experience Questionnaire Discussion

Subjects' overall agreement with 'liked the system sounds', 'enjoyed the experience', and 'system helped to get to know other subject' provides encouraging evidence that the 2Hearts system and test-tasks were engaging and pleasant. It must be noted that by the end of the experiment, each subject was on friendly terms with the experimenter, and thus these responses might have been biased. The response of (mean = 3.5, std.dev = 1.1) for 'sounds were distracting' indicates that the focus on attentional competition regarding SBFS is justified. The response of (mean = 2.3, std. dev = 1.2) for 'the sensor was distracting' indicates that the sensors employed were acceptable but might be worth improving.

The results of the freeform section responses regarding like/dislike of system audio supply useful points to be considered in future refinement of the 2Hearts audio design. The results regarding 'connection with feedback' reveal variation in how much of a connection subjects felt between their hearts and the feedback- these comments suggest some useful improvement strategies, such as allowing users to select or design a customized sound that they identify with, to represent their own heartbeat.

Users' responses regarding potential applications for 2Hearts-like systems are interesting in that they agree very well with the SBFS application categories listed in Section 1.1. It is noted that the application list presented in that section was roughly defined by the author prior to performing 2Hearts testing. Therefore these responses suggest there is a convergence of opinion regarding people's instincts about uses for SBFS's, and provide evidence that the 2Hearts system design, as a whole, conveys the intended aesthetic impressions and user experience.

12.11 Summary of Findings from User Testing

- The testing of user HR responses to audio parameter variation as employed in 2Hearts-- Version 3 gave inconclusive results. It was not determined whether the strength of pathway S→U→V is generally small or whether the system design did not effectively engage this pathway.

- The testing of users' self-rated perception of 2Hearts audio parameter variation found several significant relationships between parameter modulations and affective responses. These findings relate to pathway S→C (perhaps also S→U→C) and suggest ways to improve the 2Hearts display mapping design.

- Qualitative user responses regarding 2Hearts timbre design suggested the presence of unwanted affective bias; several users associated the pitch-sequence and noise stream with anxiety/unsettled affect and the bassline with annoying or comical affect.

- Testing of the conversation task did not find any significant relationships between HR data or questionnaire ratings and audio feedback type. This suggests that the current system design does not effectively engage pathways S→U→V or S→C. Proposed explanations are ineffective mapping design or poor match between task and system due to lack of attentional bandwidth for perceiving system audio.

- User-user HR correlation was not found to be significantly different between conversation task conditions, and was not found to be significantly greater than zero overall. However, the results do not rule out that more powerful testing could reveal such relationships, which could be used in the processing of pathways U1→V1→SS←V2←U2. It was noted that disruptive audio stimuli could cause inflated correlation measurements not necessarily arising from social interaction.

- In the control-game task, 'controller' users were able to intentionally influence the 'target' users' HR with measured significance. This implies that at least in such an extreme context, the pathway C1→E1→U2→V2 functions with notable strength. It was not shown that audio feedback enhances the intentional HR control.

13. Recommendations for Future Work

13.1 Future Improvement of the 2Hearts System

The following summarizes the most important ideas for improving the 2Hearts system that have been noted during the design process, discovered through consideration of the SBFS signal-flow model, or suggested as a result of user testing.

- Quantitative and qualitative testing results suggest that perception of the per-user low-passed/high-passed HR stream does not convey the intended dimension of arousal and does convey unwanted associations of unpleasantness/negative valence; therefore the nature of this mapping should be reconsidered. Suggested changes to address this issue include:
    - Decreasing the base frequency of the modulated pitches; this change would sacrifice some pitch perception accuracy/memorability for a reduction in perceived dissonance and an increase in suitable associations
    - Finding another scheme for reducing inter-pitch dissonance such as quantizing pitch to scale/chord tones or fixing chroma and modulating pitch-height by employing Shepard-tone style frequency-balance techniques[61]
    - Finding a non-pitch mapping for low-passed HR; for instance the display mappings for low- and high-passed HR could be swapped, and the oscillating pitch stream could be made less prominent in the audio mix
- The current filter design for separating medium-term phasic HR behaviour from HRV phenomena is not effective; more suitable digital filter designs or other frequency-isolation techniques (including predictive techniques) should be explored.
- Additional literature research into human perceptual capabilities should be undertaken; especially regarding brightness perception. This would enhance designer ability to predict perceptual resolution of SBFS audio parameters

- Several users commented that they liked hearing audio with a clear sense of beat or tempo; therefore future display-mapping designs should maintain or increase the use of measure-like looping structures.
- Several users expressed a desire for changed or customizable sounds associated with their own heartbeats; a selection or customization technique should be designed to increase users' identification with the signals and decrease negative-valence associations of the signals.
- Future design work should take advantage of heart data that has been collected to check the display of various HR phenomena and establish expected distributions for input variables.
- Future design work should take a narrower scope by focusing on 1 or 2 specific psychophysiological HR phenomena and making sure these are well-audified before adding complexity.

13.2 Improvement of the Signal-Flow Model

Recommendations for future development of the SBFS signal-flow model are:

- Given the importance of conscious versus unconscious perception in the model, continued literature research should seek to clarify the precise nature of these perceptual mechanisms
- SBFS input and output modalities other than heartbeat and audio should be examined in detail in terms of the signal-flow model. In particular, hidden-somatic input measures should be researched more given their relevance to 'essentic forms' and implicit communication.
- The composite affect-dimension suggested as 'goal-state' should be further analyzed
- It should continue to be questioned whether 'underlying internal affect dimensions' are the best portrayal of unconscious emotional state; perhaps these dimensions are more properly considered simply as the distributed collection of unconscious perceptual, visceral, ANS and somatic variables.

13.3 Improvement of Testing Procedures

Recommendations for future development of SBFS user testing procedures are:

- Several of the results hinted that a slightly larger subject sample size would yield additional significant results; therefore future tests should employ more subjects.
- Sample-listening tasks should include phases in which previous results are replicated which found a significant relationship between heart rate and music tempo/style. This would ensure that the experimental design is adequate to capture music-induced heart rate changes if they are present.
- Video/conversation/feedback data obtained from 2Hearts—Version 3 user testing should be analyzed together in detail to look for notable socially-relevant HR phenomena and to develop independent measures of social-emotional state.

- Independent measures of social-emotional state should be researched from the literature.
- Audio samples for perceptual testing should in the future be created using real recorded HR data (or realistic random data based on real HR data distribution).
- New testing strategies should be developed which directly employ the 'co-navigate' SBFS paradigm and allow users more training time, more play/free-exploration of the system and more explicit discussion of system audio.
- The 'conversation task' and 'control-game task' testing paradigms should be further refined to reduce physical-exertion based HR artifacts, reduce subjects' knowledge of experimenter-intent, and distinguish between per-user feedback path effects.

14. Conclusions

This paper has presented the results of two parallel efforts: development of a theoretical model for SBFS design and development of a particular SBFS system, 2Hearts. The theoretical exploration was begun after initial efforts on 2Hearts indicated both the promise of the SBFS area, and the difficulty of integrating SBFS design considerations without a clearly outlined description of SBFS goals and strategies.

After a cycle of building and testing both system and theory, the signal-flow model has shown its utility by allowing 2Hearts design options to be compared with clarity, by suggesting testing strategies, and by enabling testing results to be interpreted and applied to system improvement.

An important theme that has lasted throughout the 2Hearts project is the pre-existing relationship between heartbeat behaviour, musical rhythm, and emotion. It has been suggested that dynamic, multi-dimensional rhythms lay at the core of emotional state, that instinctively produced rhythmic forms may play a key role in social communication of affect, and that music itself amounts to a form of social biofeedback technology used by humans since ancient times to give expression to the dynamical, non-verbal contours of unconsciously processed affect. These time-honoured natural and artificial processes can serve as inspiration for new electronic systems to process and communicate affect through direct channels that are felt as well as understood.

The SBFS signal-flow model locates the 2Hearts system within a large space of SBFS modalities and mapping strategies. The general project of improving SBFS technology can progress through the exploration and comparison of many different regions of this design space, refining the effectiveness of display-mapping techniques for affective data. It is hoped that such endeavors will eventually allow SBFS's to convey human emotional information with a high degree of complexity and dynamism, allowing users to gain a truly new perspective on the inner rhythms of themselves and of others.

15. Summary of Contributions

The following are the most significant research contributions achieved by the present work:

1. A detailed 'signal-flow model' for describing and designing social biofeedback systems. In its current state, the model focuses on the use of heart sensor data as system input and music/audio as system output. (See Chapters 3-6 for explanation of the model and Chapter 7 for a summary of model implications).
2. A suggested high-level design procedure for SBFS's based on the signal-flow model (See Chapter 7).
3. Design and implementation details of the '2Hearts' SBFS, presented with respect to the signal-flow model and including a unique architecture for synchronized, time-compressed display streams (See Chapter 10).
4. Results from user testing of the 2Hearts system; including validation/critique of audio mapping choices, significant results suggesting that people can deliberately influence each others' heart rates through improvised non-physical communication, and qualitative findings/recommendations regarding users' experience of 2Hearts (see Chapter 12).
5. Proposal of inter-user HR correlation as a measure of social synchronization for system use and testing use; and preliminary findings indicating some drawbacks of this measure and indicating that further testing would be worthwhile (see Chapters 5, 10 and 12).
6. Three user-testing task designs ('audio-sample-listening task', 'conversation task', and 'control-game task') which could be of future use in SBFS testing (see Chapters 11 and 12)

References

[1] R. W. Picard, *Affective Computing*. Cambridge, Mass.: MIT Press, 1997.

[2] A. Strange, *Electronic Music*. Iowa: Wm. C. Brown, 1983.

[3] B. R. Knapp and H. S. Lusted, "A Bioelectric Controller for Computer Music Applications," *Computer Music Journal,* vol. 14, p. 42, 1990.

[4] A. Maebayashi (Artist). Interactive Installation: "Audible Distance." NTT Center, Hatsudai, Tokyo, 1997. Internet: http://www.ntticc.or.jp/Archive/1997/ICC_BIENNALE97/Works/audible.html [Aug. 2, 2006]

[5] C. Janney (Artist). Multimedia Dance Performance by M. Baryshnikov: "Heartbeat: mb." New York, 1998.

[6] N. Tosa (Artist). Interactive Installation: "Unconscious Flow," Siggraph technOasis, 1999. Internet: http://www.siggraph.org/artdesign/gallery/S99/artists/Tosa_Naoko.html [Aug. 2, 2006]

[7] M. Ballora, "Data Analysis through Auditory Display: Applications in Heart Rate Variability." Ph.D. thesis, McGill University, Canada, 2000.

[8] R. W. Picard and J. Scheirer, "The Galvactivator: A Glove that Senses and Communicates Skin Conductivity," in *9th International Conference on Human Computer Interaction*, 2001.

[9] J. M. DiMicco, V. Lakshmipathy, and A. T. Fiore, "Conductive Chat: Instant Messaging With a Skin Conductivity Channel (Poster)," in *Proc. Conference on Computer Supported Cooperative Work (CSCW '02)* New Orleans, LA, 2002.

[10] T. Etani and D. Tinapple (Artists). Interactive Installation: "Pimp My Heart." 2006. Internet: http://www.takehitoetani.com/hbbb.html [Aug. 2, 2006]

[11] C. Jae-woo and G. S. Vercoe, "The affective remixer: personalized music arranging," in *CHI '06 extended abstracts on Human factors in computing systems* Montréal, Québec, Canada: ACM Press, 2006.

[12] R. Hamilton, "Bioinformatic Feedback: performer bio-data as a driver for real-time composition," in *Proc. International Conference on New Interfaces for Musical Expression (NIME)*, Paris, 2006, pp. 338-341.

[13] L. R. Winer, "Biofeedback: a guide to the clinical literature," *American Journal of Orthopsychiatry,* vol. 47, pp. 626-38, 1977.

[14] P. N. Juslin, "From mimesis to catharsis: expression, perception and induction of emotion in music," in *Musical Communication*, D. Miell, R. MacDonald, and D. J. Hargreaves, Eds. New York: Oxford University Press, 2005, pp. 85-115.

[15] J. LeDoux, *The Emotional Brain.* Toronto: Simon & Schuster Paperbacks, 1996.

[16] K. Hugdahl, *Psychophysiology: The Mind-Body Perspective.* Cambridge, Mass.: Harvard University Press, 1995.

[17] S. S. Fels, "Intimacy and Embodiment: Implications for Art and Technology," in *Proc. ACM Conference on Multimedia*, 2000, pp. 13-16.

[18] Z. Pylyshyn, "The Role of Location Indexes in Spatial Perception: A Sketch of the FINST Spatial Index Model," *Cognition,* vol. 32, pp. 65-97, 1989.

[19] A. Damioso, *Descarte's error: Emotion, reason, and the human brain.* New York: Grosset/Putnam, 1994.

[20] J. Sweller, "Cognitive Load Theory, learning difficulty, and instructional design," *Learning and Instruction,* vol. 4, pp. 295-312, 1994.

[21] A. Ehlers and J. Margraf, "Anxiety induced by false heart rate feedback in patients with panic disorder," *Behaviour Research and Therapy,* vol. 26, pp. 1-11, 1987.

[22] S. Valins, "Cognitive effects of false heart-rate feedback.," *Journal of Personality and Social Psychology,* vol. 4, pp. 400-408, 1996.

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

[24] J. A. Sloboda and P. N. Juslin, "Psychological Perspectives on Music and Emotion," in *Music and Emotion: Theory and Research*, P. N. Juslin and J. A. Sloboda, Eds. New York: Oxford University Press, Inc., 2001, pp. 71-104.

[25] K. Oatley, *Best laid schemes. The psychology of emotions.* Cambridge, MA: Harvard University Press, 1992.

[26] R. S. Lazarus, *Emotion and Adaptation.* New York: Oxford University Press, 1991.

[27] P. Ekman, "An argument for basic emotions," *Cognition and Emotion,* vol. 6, pp. 169-200, 1992.

[28] Task Force of the European Society of Cardiology, North American Society for Pacing and Electrophysiology, "Heart rate variability – standards of measurement, physiological interpretation, and clinical use," *Circulation,* vol. 93, pp. 1043-1065, March 1 1996.

[29] T. Thong, K. Li, J. McNames, M. Aboy, and B. Goldstein, "Accuracy of ultra-short term heart rate variability measures," in *Proc. IEEE EMBS Conference,* Cancun, Mexico, 2003, pp. 2424-2427.

[30] N. Hayes, *Foundations of Psychology*. New York: Routledge, 1994.

[31] P. A. Obrist, "The cardiovascular-behavioral interaction as it appears today," *Psychophysiology,* vol. 13, pp. 95-107, 1976.

[32] D. C. Fowles, "Psychophysiology and psychopathology: A motivational apprach," *Psychophysiology,* vol. 25, 1988.

[33] Venables, "Autonomic activity," *Annals of the New York Academy of Sciences,* vol. 620, pp. 191-207, 1991.

[34] J. I. Lacey, "Somatic response patterning and stress: Some revisions of activation theory," in *Psychological stress: Issues in research*, M. H. Appley and R. Trumbull, Eds. New York: Appleton-Century-Crofts, 1967.

[35] S. W. Porges, "Respiratory sinus arrhythmia: Physiological basis, quantitative methods, and clinical implications," in *Cardiorespiratory and cardiosomatic psychophysiology*, P. Grossman, K. Janssens, and D. Vaitl, Eds. New York: Plenum Press, 1986, pp. 101-115.

[36] P. Ekman, R. W. Levenson, and W. V. Friesen, "Autonomic nervous system activity distinguishes among emotions," *Science,* vol. 221, pp. 1208-1210, 1983.

[37] C. Collet, E. Vernet-Maury, G. Delhomme, and A. Dittmar, "Autonomic Nervous System Response Patterns Specificity to Basic Emotions," *Journal of the Autonomic Nervous System,* vol. 62, pp. 45-57, 1997.

[38] R. W. Picard, J. Healey, and E. Vyzas, "Toward Machine Emotional Intelligence Analysis of Affective Physiological State," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 23, pp. 1175-1191, 2001.

[39] A. J. Lees, "Facial mannerisms and tics," *Advances in Neurology,* vol. 49, pp. 255-261, 1988.

[40] P. Ekman, "Should We Call it Expression or Communication?," *Innovations in Social Science Research,* vol. 10, pp. 333-344, 1997.

[41] R. J. R. Blair, "Facial expressions, their communicatory functions and neuro-cognitive substrates," *Philosophical Transactions of the Royal Society of London*, vol. B, pp. 561-572, 2003.

[42] U. Dimberg, M. Thunberg, and K. Elmehed, "Unconscious facial reactions to emotional facial expressions," *Psychological Science*, vol. 11, pp. 86-89, 2000.

[43] P. J. Whalen, S. L. Rauch, N. L. Etcoff, S. C. McInerney, M. B. Lee, and M. A. Jenike, "Masked Presentations of Emotional Facial Expressions Modulate Amygdala Activity without Explicit Knowledge," *Journal of Neuroscience*, vol. 18, pp. 411-418, January 1 1998.

[44] L. M. Williams, B. J. Liddell, A. H. Kemp, R. A. Bryant, R. A. Meares, A. S. Peduto, and E. Gordon, "Amygdala-prefrontal dissociation of subliminal and supraliminal fear," *Human Brain Mapping*, November 9 2005.

[45] M. Clynes, *Sentics: The touch of emotions*. New York: Doubleday, 1977.

[46] S. Baron-Cohen and S. Wheelwright, "The Empathy Quotient (EQ). An investigation of adults with Asperger Syndrome or High Functioning Autism, and normal sex differences," *Journal of Autism and Developmental Disorders*, vol. 34, pp. 163-175, 2004.

[47] R. W. Doherty, "The Emotion Contagion Scale: A measure of individual differences," *Journal of Nonverbal Behavior*, vol. 21, pp. 131-154, 1997.

[48] S. Whittaker, "Theories and Methods in Mediated Communication," in *The Handbook of Discourse Processes*, A. Graesser, M. Gernbacher, and S. Goldman, Eds.: Erlbaum, 2002, pp. 243-286.

[49] C. D. Marci and S. P. Orr, "The Effect of Emotional Distance on Psychophysiologic Concordance and Perceived Empathy Between Patient and Interviewer," *Applied Psychophysiology and Biofeedback*, May 25 2006.

[50] P. V. Leeuwen, D. Geue, S. Lange, D. Cysarz, H. Bettermann, and D. H. W. Grönemeyer, "Is there evidence of fetal-maternal heart rate synchronization?," in *BioMed Central Physiology*. vol. 3, 2003.

[51] V. S. Anishchenko, A. G. Balanov, N. B. Janson, N. B. Igosheva, and G. V. Bordyugov, "Entrainment between heart rate and weak noninvasive forcing," *Internation Journal of Bifurcation & Chaos*, vol. 10, pp. 2339-2348, 2000.

[52] S. M. Pincus, T. Mulligan, A. Iranmanesh, S. Gheorghiu, M. Godschalk, and J. D. Veldhuis, "Older males secrete luteinizing hormone and testosterone more irregularly, and jointly more asynchronously, than younger males," *Proc.*

National Academy of Sciences of the U.S.A., vol. 93, pp. 14100-14105, November 1996.

[53] B. Pompe, P. Blidh, D. Hoyer, and M. Eiselt, "Using mutual information to measure coupling in the cardiorespiratory system," *IEEE Engineering in Medicine and Biology Magazine*, vol. 17, pp. 32-39, Nov-Dec 1998.

[54] G. Kramer, "An Introduction to Auditory Display," in *Auditory Display: Sonification, Audification, and Auditory Interfaces*, G. Kramer, Ed. Reading, MA: Addison Wesley, 1994, pp. 1-77.

[55] M. Matthews, "The Ear and How It Works," in *Music, cognition, and computerized sound: an introduction to psychoacoustics*, P. R. Cook, Ed. Cambridge, Mass: MIT Press, 1999, pp. 1-10.

[56] M. Matthews, "The Auditory Brain," in *Music, cognition, and computerized sound: an introduction to psychoacoustics*, P. R. Cook, Ed. Cambridge, Mass: MIT Press, 1999, pp. 11-20.

[57] M. Matthews, "What Is Loudness?," in *Music, cognition, and computerized sound: an introduction to psychoacoustics*, P. R. Cook, Ed. Cambridge, Mass: MIT Press, 1999, pp. 71-78.

[58] J. Pierce, "Hearing in Time and Space," in *Music, cognition, and computerized sound: an introduction to psychoacoustics*, P. R. Cook, Ed. Cambridge, Mass: MIT Press, 1999, pp. 89-104.

[59] J. Pierce, "Introduction to Pitch Perception," in *Music, cognition, and computerized sound: an introduction to psychoacoustics*, P. R. Cook, Ed. Cambridge, Mass: MIT Press, 1999, pp. 57-70.

[60] J. Pierce, "Consonance and Scales," in *Music, cognition, and computerized sound: an introduction to psychoacoustics*, P. R. Cook, Ed. Cambridge, Mass: MIT Press, 1999, pp. 167-185.

[61] R. Shepard, "Pitch Perception and Measurement," in *Music, cognition, and computerized sound: an introduction to psychoacoustics*, P. R. Cook, Ed. Cambridge, Mass: MIT Press, 1999, pp. 149-166.

[62] D. J. Levitin, "Memory for Musical Attributes," in *Music, cognition, and computerized sound: an introduction to psychoacoustics*, P. R. Cook, Ed. Cambridge, Mass: MIT Press, 1999, pp. 209-228.

[63] G. Kramer, "Some Organizing Principles for Representing Data with Sound," in *Auditory Display: Sonification, Audification, and Auditory Interfaces*, G. Kramer, Ed. Reading, MA: Addison Wesley, 1994, pp. 195-221.

[64] J. M. Grey, "Multidimensional Perceptual Scaling of Musical Timbres," *Journal of the Acoustical Society of America,* vol. 61, pp. 1270-1277, 1976.

[65] R. Bargar, "Pattern and Reference in Auditory Display," in *Auditory Display: Sonification, Audification, and Auditory Interfaces*, G. Kramer, Ed. Reading, MA: Addison Wesley, 1994, pp. 151-165.

[66] T. McLaughlin, *Music and Communication*. London: Faber and Faber, 1970.

[67] A. Schoenberg, *Theory of Harmony*, 3 ed. Los Angeles: University of California Press, 1922.

[68] L. B. Meyer, *Emotion and Meaning in Music*. Chicago: Chicago University Press, 1956.

[69] R. Shepard, "Cognitive Psychology and Music," in *Music, cognition, and computerized sound: an introduction to psychoacoustics*, P. R. Cook, Ed. Cambridge, Mass: MIT Press, 1999, pp. 21-36.

[70] S. M. Williams, "Perceptual Principles in Sound Grouping," in *Auditory Display: Sonification, Audification, and Auditory Interfaces*, G. Kramer, Ed. Reading, MA: Addison Wesley, 1994, pp. 95-125.

[71] W. F. Thompson and B. Robitaille, "Can composers express emotions through music?," *Emprical Studies of the Arts,* vol. 10, pp. 79-89, 1992.

[72] 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, 2003.

[73] K. R. Scherer and M. R. Zentner, "Emotional Effects of Music: Production Rules," in *Music and Emotion: Theory and Research*, P. N. Juslin and J. A. Sloboda, Eds. New York: Oxford University Press, 2001, pp. 361-392.

[74] M. R. Zentner and J. Kagan, "Infants' perception of consonance and dissonance in music," *Infant Behavior and Development,* vol. 21, pp. 483-492, 1998.

[75] H. M. Borchgrevink, "Musikalske akkod-prefereanser hos mennesket belyst ved dyreforsok [Musical chord preferences in humans as demonstrated through animal experiments]," *Tidskrift for den Norske Laegefoerning,* vol. 95, pp. 356-358, 1975.

[76] B. DeVries, "Assessment of the affective response to music with Clynes' sentograph," *Psychology of Music,* vol. 19, pp. 46-64, 1991.

[77] J. A. Sloboda and P. N. Juslin, "Music and Emotion: Commentary," in *Music and Emotion: Theory and Research*, P. N. Juslin and J. A. Sloboda, Eds. New York: Oxford University Press, Inc., 2001, pp. 453-462.

[78] D. L. Bartlett, "Physiological responses to music and sound stimuli," in *Handbook of music psychology*, 2 ed, D. A. Hodges, Ed. San Antonio, TX: IMR, 1996, pp. 343-385.

[79] S. Khalfa, I. Peretz, and J.-P. Blondin, "Event-related skin conductance responses to musical emotions in humans," *Neuroscience Letters,* vol. 328, pp. 145-149, August 9 2002.

[80] G. Harrer and H. Harrer, "Music, Emotion and Autonomic Function," in *Music and the Brain: Studies in the Neurology of Music*, M. Critchley and R. A. Henson, Eds., 1977.

[81] H. Katayose, S. Hirai, T. Kanamori, H. Kato, and S. Inokuchi, "Physiological Measurement of Performer's Tension and its Utilization for Media Control," in *Proc. International Computer Music Conference (ICMC)*, 1999.

[82] L. Bernardi, C. Porta, and P. Sleight, "Cardiovascular, cerebrovascular, and respiratory changes induced by different types of music in musicians and non-musicians: the importance of silence.," *Heart,* vol. 92, pp. 433-434, 2006.

[83] C. L. Darner, "Sound pulses and the heart," *Journal of the Acoustical Society of America,* vol. 39, pp. 414-416, 1966.

[84] D. S. Ellis and C. Brighouse, "Effects of music on respiration- and heart-rate," *American Journal of Psychology,* vol. 65, pp. 414-416, 1952.

[85] "Vernier Software & Technology Home Page." Internet: http://www.vernier.com [July 18, 2006].

[86] "Thought Technology, Inc. Home Page." Internet: http://www.thoughttechnology.com/index.htm [Aug. 2, 2006]

[87] "Polar Inc. Home Page." Internet: http://www.polar.fi [July 18, 2006]

[88] T. Marrin and R. Picard, "The Conductor's Jacket: a Device for Recording Expressive Musical Gestures.," in *Proc. International Computer Music Conference* Ann Arbor, MI, 1998, pp. 215-219.

[89] C. Peter, E. Ebert, and H. Beikirch, "A wearable multi-sensor system for mobile acquisition of emotion-related physiological data.," in *Proc. First International Conference on Affective Computing and Intelligent Interaction: ACII 2005*, 2005, pp. 691-698.

[90] T. M. Nakra, "Inside the Conductor's Jacket: Analysis, Interpretation and Musical Synthesis of Expressive Gesture." Ph.D. thesis, Massachusetts Institute of Technology, U.S.A., 2000. Available: http://vismod.media.mit.edu/tech-reports/TR-518/ [Aug. 2, 2006]

[91] "Wireless 2000- Heart & Respiration Rate Monitor." Internet: http://www.wireless200.com/hrrm.htm [July 18, 2006]

[92] F. Déchelle, "jMax : un environnement pour la réalisation d'applications musicales sur Linux," in *Journées d'Informatique musicale* Bordeaux, 2000.

[93] M. Puckette, "Pure Data," in *Proc. International Computer Music Conference (ICMC)*, San Francisco, 1996, pp. 269-272.

[94] R. G. McCaig and S. S. Fels, "Playing on Heart-Strings: Experiences with the 2Hearts System," in *Proc. 2nd International Conference on New Interfaces for Musical Expression (NIME02)*, Dublin, 2002, pp. 54-59.

[95] A. Bierbaum, C. Just, P. Hartling, K. Meinert, A. Baker, and C. Cruz-Neira, "VR Juggler: A virtual platform for virtual reality application development," in *IEEE VR*, Yokohama, Japan, 2001, pp. 89-96.

[96] "Matlab Central Exchange Website." Internet: http://www.mathworks.com/matlabcentral/fileexchange/ [Aug. 2, 2006]

Appendix A. Detailed Functional Description of 2Hearts Software Implementation

This section explains the operation of the 2Hearts-- Version 3 software by describing the inputs/outputs and functionality of each subpatch in turn (refer to Figure A-1 below).

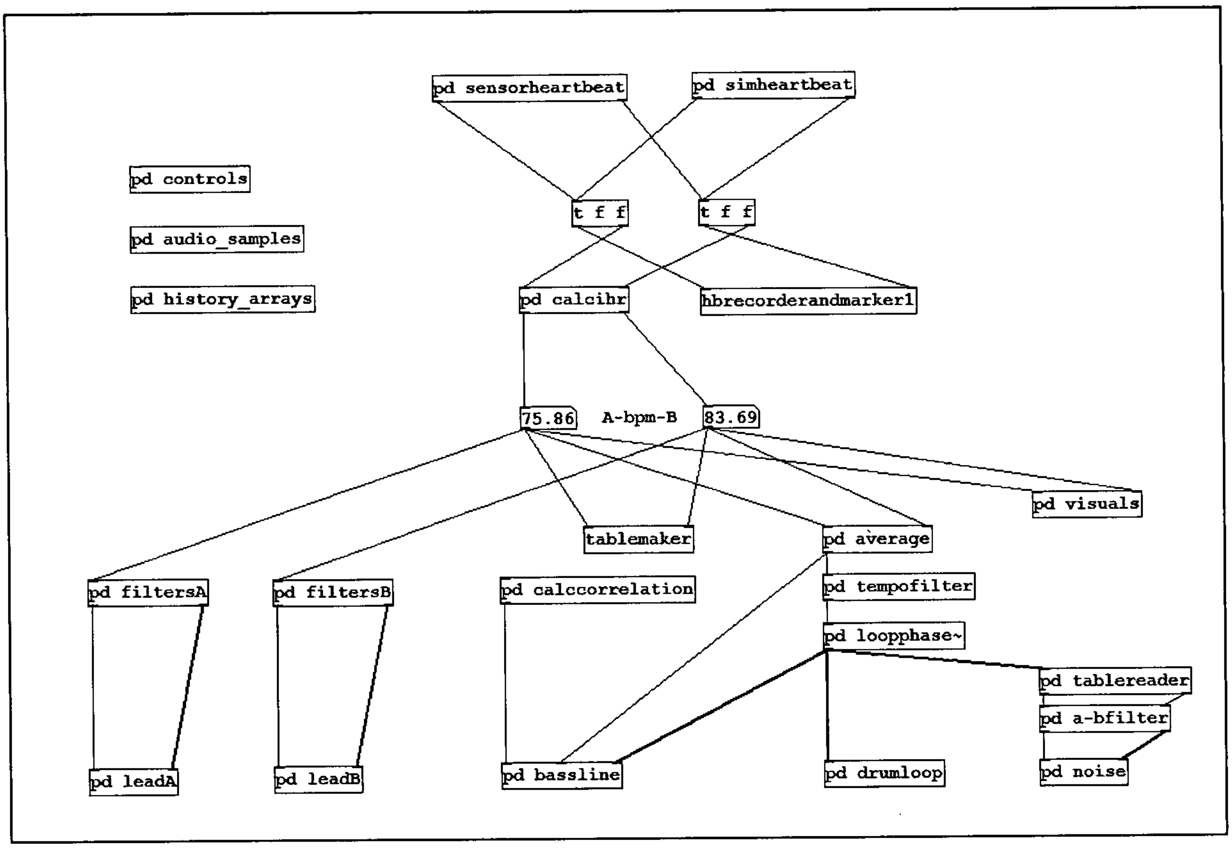


Figure A-1. Main PD patch window of 2Hearts--Version 3. Audio output modules are grouped along the bottom row, and heart data input sources along the top. The boxes containing "t f f" are PD 'trigger' routing objects which pass incoming data to both outputs, in right-output then left-output order. Note: Stored sample and array data can be accessed by named reference within the PD environment without the need for explicit signal connections.

Many of the filter and range-mapping constants employed in the 2Hearts-Version 3 software were determined empirically by testing various parameter values with the designer's own heart data and with a file of 10-minute sample data recorded from two volunteers.

Subpatch 'sensorheartbeat'

Outputs: Sends trigger signals for detected user1/user2 heartbeats to 'calcihr'.

Function: This subpatch internally reads audio data from the A/D converter and performs peak-detection to create heartbeat-trigger signals within the PD patch. Peak-detection is performed by a custom 'heartbeat~' code module added to the PD framework (see Appendix B for code listing). This module performs peak-detection on each heart-sensor signal using a high/low threshold technique. Appropriate high thresholds for the two heart monitors were found to be 0.01 and 0.001 of maximum audio amplitude. Timing precision of the peak-detected triggers is one PD DSP tick (although the 'heartbeat~' code is capable of detecting triggers with per-sample precision). A simple error-rejection technique cancels triggers falling within 300ms of the previous triggers.

Subpatch 'simheartbeat'

Outputs: Sends randomly generated heartbeat-triggers to 'calcihr'.

Function: This subpatch was used for system development and to generate sample audio excerpts for user testing. It creates 2 streams of random trigger data by selecting a new IBI value from an evenly distributed probability interval after each trigger sent (different minimum/maximum IBI values can be set for each of the 2 streams).

Subpatch 'calcihr'

Inputs: Receives heartbeat-trigger signals from 'sensorheartbeat' or 'simheartbeat'.

Outputs: Sends IHR values (in BPM) from each of the two heart-data streams to many of the other subpatches; values are sent once per detected heartbeat.

Function: For each user-data stream, for each incoming trigger signal, the time interval is measured between the current time and the previous trigger time. This interval is converted into a rate in BPM.

Subpatch 'hbrecorderandmarker1'

Inputs: Receives heartbeat-trigger signals from 'sensorheartbeat' or 'simheartbeat'.

Function: This subpatch records heartbeat trigger time series data, measured in PD DSP ticks, into arrays, which can be saved to ASCII text files. It also records and saves time-markers for user-testing events such as trial-start and trial-end, into separate files as paired time/marker-type values.

Subpatch 'tablemaker'

Inputs: IHR values from 'calcihr'.

Function: This subpatch samples the last-received IHR values once per DSP tick into a circular array of length 1000, for each user-data stream.

Subpatch 'average'

Inputs: IHR values from 'calcihr'.

Outputs: Mean IHR value, sent to 'tempofilter' and to 'bassline'.

Function: For each new incoming IHR value from either user, calculates the mean of the 2 users' most recent IHR values.

<u>Subpatch 'tempofilter'</u>

Inputs: Mean(IHR1,IHR2) values from subpatch 'average'.

Outputs: Mean IHR values, smoothed to prevent quick variation, to subpatch 'loopphase~'.

Function: This subpatch receives mean IHR values which change stepwise for each new heartbeat, and filters these values to produce a smoothly varying signal. The built-in PD filter object 'lop~' (a one-pole lowpass filter) is used with a rolloff frequency of 0.01Hz.

<u>Subpatch 'loopphase~'</u>

Inputs: Smoothed mean IHR signal from 'tempofilter'.

Outputs: To 'drumloop~', 'tablereader' and 'bassline', a modulated ramp signal from 0.0 to 1.0 indicating phase progress through the sliding/looping window.

Function: This subpatch creates a phase signal used to drive the bassline, drum and noise loops in synchronization. Progress through a 16-beat loop occurs at a rate controlled by smoothed mean HR, according to the formula: Looping Frequency (in loops/s) = Mean HR (in BPM) / 960 (the constant equal to 16 beats * 60 s/min). The PD 'phasor~' object is used to generate monotonically increasing values from 0.0 to 1.0 based on this looping frequency.

<u>Subpatches 'filtersA'/ 'filtersB'</u>

Inputs: IHR values from 'calcihr'.

Outputs: Low- and high- pass filtered IHR signals to 'leadA'/ 'leadB'.

Function: One copy of the IHR input stream, in BPM, is processed by the PD filter 'lop~' with rolloff frequency of 0.02Hz. Another copy of the input stream is multiplied by 20, then increased by 100 BPM, then processed by the PD filters 'hip~' and 'lop~' in series, with rolloffs of 0.05Hz and 20Hz respectively. Low-passed values are sampled and sent once per incoming new IHR value.

<u>Subpatches 'leadA'/ 'leadB'</u>

Inputs: Low- and high-pass filtered IHR signals from 'filtersA'/ 'filtersB'.

Function: This subpatch produces audio for the 2Hearts per-user HR display streams, using a sample playback engine based on PD's "chocolate sampler" demonstration patch. The sample is played from beginning once per incoming low-passed IHR value. Sample pitch is calculated as: Pitch (in Hz) = Lowpassed IHR (in BPM) x 8 Hz/BPM. Sample audio is processed with 2 'lop~' filter objects in series. Rolloff freq. for each of these filters is calculated as: Filter Freq. (in Hz) = (Highpassed HR (in BPM) * 1 Hz/BPM) + 200Hz). Filter Freq. is limited with a lower threshold of 120 Hz.

<u>Subpatch 'calcorrelation'</u>

Outputs: Non-normalized inter-user HR correlation values to 'bassline'.

Function: Ten times per second, this subpatch reads from the circular past-IHR arrays to calculate a correlation value similar to the Pearson's correlation (except with no normalizing dividend according to total signal power). This correlation measure ('Correlation2H') is calculated as:

$$\textbf{Correlation2H} = \sum [(\mathbf{x_n} - \text{mean}(\mathbf{x})) * (\mathbf{y_n} - \text{mean}(\mathbf{y}))]$$

Where $\mathbf{x_n}$, $\mathbf{y_n}$ are stored IHR array values of index **n** from user 1 and user 2, respectively, in BPM, mean(**x**) and mean(**y**) are the mean values of the total user 1/user2 arrays, respectively, in BPM, and the sum is calculated for array indices **n** = 1 to 1000. The lack of normalization has two consequences: Correlation2H varies with array length, which is of no concern since the length is fixed; and Correlation2H gives more extreme values when input signals vary greatly from their means. The latter property was considered acceptable since if both users are undergoing dramatic HR changes but in a synchronized fashion, this may indicate a greater level of empathy than small HR changes that happen in synch.

<u>Subpatch 'bassline'</u>

Inputs: Correlation signal from 'calcorrelation', mean IHR signal from 'average' and phase signal from 'loopphase~'.

Function: This subpatch produces audio for the 2Hearts correlation display stream. The audio stream is initiated from a sawtooth oscillator (the PD 'phasor~' object). Oscillator pitch is determined by mean IHR input according to: pitch (in Hz) = Mean HR (in BPM) * 0.6 Hz/BPM. The audio is processed by PD's 'vcf~' bandpass variable-q filter with q set to 3 and centre-frequency determined by: filter freq. (in Hz) = (Correlation2H * -0.015 Hz) + 300 Hz). Centre-frequency is limited by a lower threshold of 50 Hz. Audio is then additionally processed by 'lop~' and 'hip~' objects with rolloff freq.'s of 600Hz and 5Hz, respectively. Audio amplitude is then modulated by values read from a table of length 1000 which is traversed once per loop phase; table values were drawn in arbitrarily to add a small amount of natural-sounding volume variation to the bassline and encourage perception of a common looping period between bassline, noise and drum streams.

Subpatch 'drumloop~'

Inputs: Loop phase value from 'loopphase~'.

Function: This subpatch plays back a sampled audio loop of 16 identical regularly-timed synthesized drum beats. The audio loop is traversed at a variable rate according to loop phase data, such that a mean HR of *X* BPM will lead to drum beats being played at the same rate, *X* BPM.

Subpatch 'tablereader'

Inputs: Loop phase value from 'loopphase~'.

Outputs: Interpolated IHR data from the stored IHR arrays, to 'a-bfilter'.

Function: This subpatch calculates an interpolated value from the arrays of past IHR data, based on the array position given by the phase input value (where 0 gives array start and 1 gives array end). Four-point interpolation is employed using PD's 'tabread4~' object.

Subpatch 'a-bfilter'

Inputs: Stored windowed IHR data from 'tablereader'.

Outputs: Two differently filtered streams to 'noise', based on the difference between IHR1 and IHR2.

Function: Each incoming IHR stream is filtered through 'hip~' (with rolloff frequency of 0.1Hz), then the absolute value of the difference in IHR values is calculated. The first output sends this difference. The difference is filtered further with 'lop~' (with rolloff frequency of 0.01Hz) for transmission from the second output.

Subpatch 'noise'

Inputs: Two streams of IHR difference data from 'a-bfilter'.

Function: This subpatch creates the 2Hearts audio display stream for user-user HR difference. Audio is initiated with PD's 'noise~' object, and filtered by PD's 'vcf~' object with q set to 3. Filter centre-frequency is modulated by the first difference-input stream. Before controlling the filter, the input stream is multiplied by 2, increased by 70 BPM, scaled logarithmically through PD's 'tabread mtof' object, and filtered by 'lop~' with rolloff freq. of 1.5Hz. For timbre design, audio is additionally processed by a 'lop~' and two 'hip~' objects in series, with rolloff values of 700Hz, 4kHz and 2.5kHz respectively. The audio amplitude is scaled according to the second, smoothed input stream; according to: amplitude-scaling-factor = (IHRdifference (in BPM))^2 * 0.00025 BPM^{-2} + 0.1. Amplitude factor is limited to a maximum value of 7.

Subpatch 'visuals'

Inputs: IHR values from 'calcihr'.

Function: When double-clicked, this subpatch opens a window showing a simple visual display of each user's IHR, employing PD's built-in 'vumeter' object. A linear vertical scale is shown corresponding to a range of 50 BPM to 130 BPM. Figure A-2 shows the display created by 'visuals'.

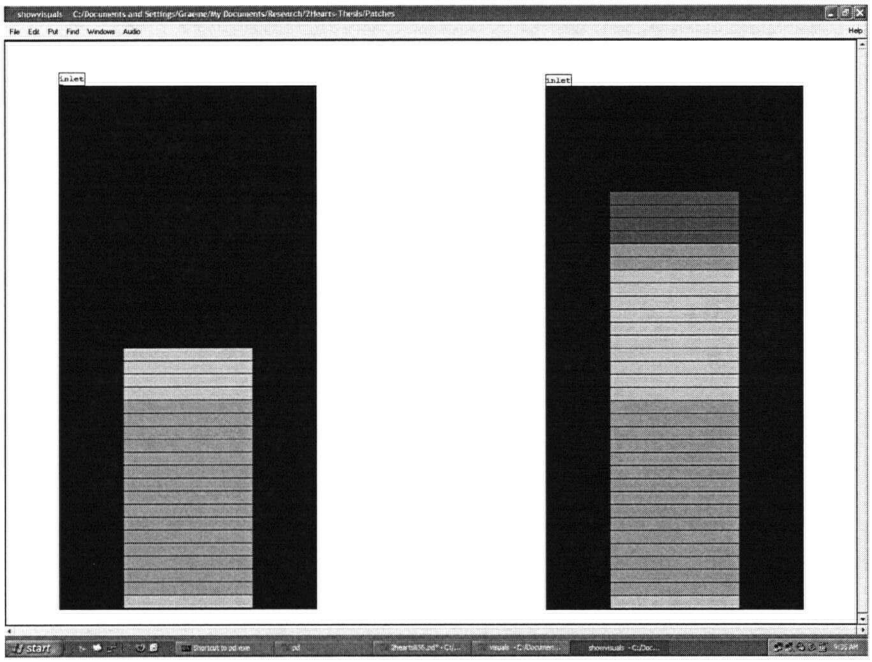


Figure A-2. Heart Rate Display of 2Hearts 'Visuals' Object.

Subpatch 'controls'

Function: When double-clicked, this subpatch opens a window showing various operator controls for the 2Hearts software. Controls are included for enabling/disabling simulated and real heart data input, saving/clearing stored heart data, and creating time markers during experimentation. A set of mixer volume controls allows the volume of each audio stream to be modified. Figure A-3 below shows the 'controls' window.

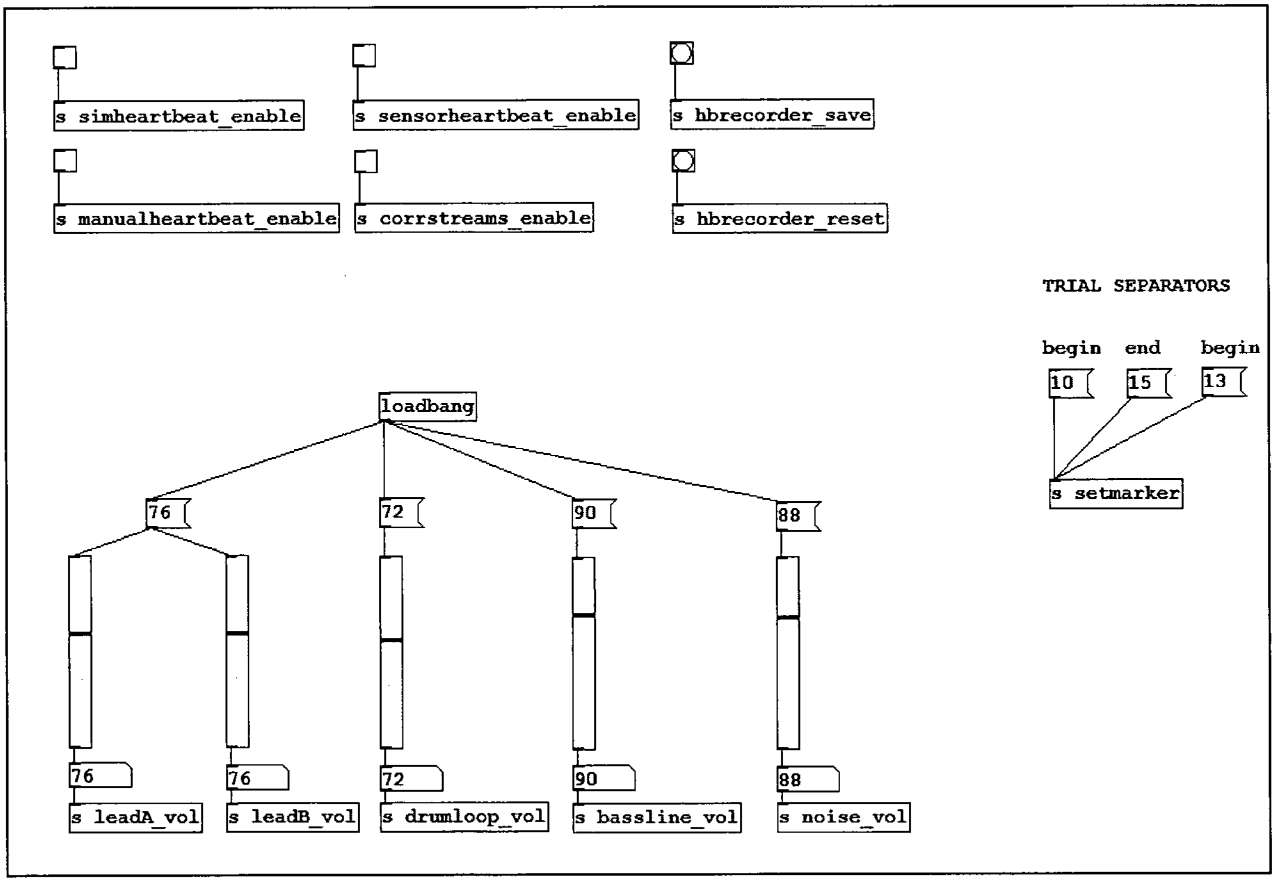


Figure A-3. Experimenter Controls Window of 2Hearts--Version 3.

Not described here are 2 capabilities left out of the 2Hearts--Version 3 patch employed in user testing: manual heartbeat triggering and playback of stored heart data. These capabilities were built into development versions of the system for use in the design process. The heart-data playback module produced heartbeat triggers according to the files saved by 'hbrecorderandmarker1', with a precision of 1 DSP tick.

As discussed, several of the constants employed in the 2Hearts--Version 3 PD patch were chosen by heuristic/qualitative methods. For free parameters relating to timbre design, this procedure may continue to play an important part of 2Hearts (and other SBFS) development; good timbre design requires a degree of artistry (as well as the incorporation of frequent test-user feedback).

For the constants relating to range mapping/warping, the SBFS signal-flow model should be employed in the future to create a more structured design procedure. Real user heart data should be employed to test for optimal perception of known HR patterns; and distribution statistics of real heart data should be used to set range-mapping values. The use of digital filter designs other than the built-in PD filters should also be explored for more effective isolation of particular signal phenomena.

Appendix B. Code Listing for Heartbeat Input Object

Note: The symbol "►" is used to indicate points where single lines of code are 'wrapped around' to the next printed line.

```
#include "m_pd.h"
#ifdef NT
#pragma warning( disable : 4244 )
#pragma warning( disable : 4305 )
#endif

/* ------------------------ heartbeat~ ----------------------------- */
/* signal object to do peak-detection and timestamping of heartbeats.*/
/* Author: R. Graeme McCaig   Last Modified: July 31, 2006 */
/* Based on PD example file dspobj~.c */


static t_class *heartbeat_class;

typedef struct _heartbeat
{
    t_object x_obj;     /* obligatory header */
    t_outlet *x_out2;
    t_outlet *x_out3;
    t_outlet *x_out4;
    t_outlet *x_out5;
    t_float x_f;        /* place to hold inlet's value if it's set by
►message */
    t_float x_hit1; /* hi threshold */
    t_float x_hit2;
    t_float x_lot1; /* lo threshold */
    t_float x_lot2;
    t_int x_tickcount;  /* number of dsp ticks since heartbeat object
►was created, used for timestamping */
    t_int x_detectstate; /* holds state for thresholding peak-detection
►method */
    t_int x_detectedonsample;
    t_int x_detectstate2; /* holds state for thresholding peak-
►detection method */
    t_int x_detectedonsample2;
} t_heartbeat;

    /* this is the actual performance routine which acts on the
►samples.
    It's called with a single pointer "w" which is our location in the
    DSP call list.  We return a new "w" which will point to the next
►item
    after us.  Meanwhile, w[0] is just a pointer to dsp-perform itself
    (no use to us), w[1] and w[2] the input vector locations,
    w[3] should be the pointer to x,
    and w[4] is the number of points to calculate. */
```

```
static t_int *heartbeat_perform(t_int *w)
{
    t_float *in = (t_float *)(w[1]);
    t_float *in2 = (t_float *)(w[2]);

    t_heartbeat *x = (t_heartbeat *)(w[3]);

    t_float hit1 = x->x_hit1;
    t_float hit2 = x->x_hit2;
    t_float lot1 = x->x_lot1;
    t_float lot2 = x->x_lot2;

    int n = (int)(w[4]);

    int detected = 0;
    int detected2 = 0;

    int i;

    x->x_tickcount = x->x_tickcount + 1;


    if (!(x->x_detectstate))
    {
      for (i = 0; i < n; i++)
      {
         if (in[i] > hit1) {detected = 1; x->x_detectstate = 1; x-
►>x_detectedonsample = i; break;}

      }
    }
    else
    {
      for (i = 0; i < n; i++)
      {
         if (in[i] < lot1) {x->x_detectstate = 0; break;}

      }
    }


    if (!(x->x_detectstate2))
    {

      for (i = 0; i < n; i++)
      {
         if (in2[i] > hit2) {detected2 = 1; x->x_detectstate2 = 1; x-
►>x_detectedonsample2 = i; break;}

      }
    }
    else
    {
      for (i = 0; i < n; i++)
      {
```

```
            if (in2[i] < lot2) {x->x_detectstate2 = 0; break;}

        }
    }

    if (detected) {outlet_float(x->x_out3, n); outlet_float(x->x_out2,
▶x->x_detectedonsample); outlet_float(x->x_obj.ob_outlet, x-
▶>x_tickcount);}
    if (detected2) {outlet_float(x->x_out5, x->x_tickcount);
▶outlet_float(x->x_out4, x->x_detectedonsample2);}

    return (w+5);
}

    /* called to start DSP.  Here we call Pd back to add our perform
    routine to a linear callback list which Pd in turn calls to grind
    out the samples. */

static void heartbeat_dsp(t_heartbeat *x, t_signal **sp)
{
    dsp_add(heartbeat_perform, 4, sp[0]->s_vec, sp[1]->s_vec, x, sp[0]-
▶>s_n);
}


    /* called to instantiate each instance of the heartbeat object */


static void *heartbeat_new(t_floatarg thresh1a, t_floatarg thresh1b,
t_floatarg thresh2a, t_floatarg thresh2b)
{
    t_heartbeat *x = (t_heartbeat *)pd_new(heartbeat_class);
    inlet_new(&x->x_obj, &x->x_obj.ob_pd, &s_signal, &s_signal);
    outlet_new(&x->x_obj, &s_float);
    x->x_out2 = outlet_new(&x->x_obj, &s_float);
    x->x_out3 = outlet_new(&x->x_obj, &s_float);
    x->x_out4 = outlet_new(&x->x_obj, &s_float);
    x->x_out5 = outlet_new(&x->x_obj, &s_float);
    x->x_f = 0;
    x->x_tickcount = 0;
    x->x_detectstate = 0;
    x->x_detectedonsample = 0;
    x->x_detectstate2 = 0;
    x->x_detectedonsample2 = 0;
    x->x_hit1 = thresh1a;
    x->x_hit2 = thresh1b;
    x->x_lot1 = thresh2a;
    x->x_lot2 = thresh2b;
    return (x);
}

    /* this routine, which must have exactly this name (with the "~"
▶replaced
    by "_tilde) is called when the code is first loaded, and tells Pd
▶how
    to build the "class". */
```

```
void heartbeat_tilde_setup(void)
{
    heartbeat_class = class_new(gensym("heartbeat~"),
►(t_newmethod)heartbeat_new, 0,
►     sizeof(t_heartbeat), 0, A_DEFFLOAT, A_DEFFLOAT, A_DEFFLOAT,
►A_DEFFLOAT, 0);
          /* this is magic to declare that the leftmost, "main" inlet
          takes signals; other signal inlets are done differently... */
    CLASS_MAINSIGNALIN(heartbeat_class, t_heartbeat, x_f);
      /* here we tell Pd about the "dsp" method, which is called back
      when DSP is turned on. */
    class_addmethod(heartbeat_class, (t_method)heartbeat_dsp,
gensym("dsp"), 0);
}
```

Appendix C. UBC BREB Ethics Approval Certificate

The ethics form is reproduced on the next page.

Appendix D. Copy of Consent Form / Website

The consent form begins on the next page.

The website contained identical text to the consent form.

Human Communication Technologies Lab
Department of Electrical & Computer Engineering
2356 Main Mall, Room 155A
Vancouver, BC, Canada, V6T 1Z4

Consent Form:

User Testing of the 2Hearts Social Heartbeat Display System

This study will test the 2Hearts music system, which produces background music for two people as they engage in social interaction. You and another subject will each be wearing a heart rate monitor, which fits around the chest, against the skin, and measures your heart rate. In the preliminary phase of the experiment, you will be asked to rest silently for a short period, and then listen to several musical excerpts. In the main phase of the experiment, you and the other subject will be instructed at various times to converse normally, to not talk, or to talk about specific topics given by cards from a board game. The system will generate music which varies in character depending on the activity of the heart beat signals. At times the sound will be turned off, and at times the system will be set to play sounds not related to your heart beat signals. In the final phase of the experiment, we will ask you to listen to several more excerpts of music, and rate them according to how you feel.

We will videotape you and record your heartbeat signals, to analyze how the system affects the social interaction and to determine how it can be improved in future versions. We will also ask you to fill out a questionnaire describing your experience of using the system. Your participation will last 2 hours or less, taking place in one session, and you will be paid an honorarium of $10 for your participation. If you have any questions or inquiries about the experimental procedures, please ask and we will provide additional explanation.

We will ensure that the recordings are kept secure in a locked faculty office. Only the experimenters will be allowed access to the data. All data from individual subjects will be coded in any publicly available reports, articles and presentations that result from this work, so that your identity is kept confidential. Videos from the study may be publicly presented; in which case we will employ video editing techniques to ensure your identity is completely obscured. This study is being conducted as part of graduate thesis work, and results will be published in the thesis.

We intend for your experience in this study to be pleasant and not stressful in any way. **If you are uncomfortable or unhappy participating in the study you may withdraw at any time**. We will be pleased to explain the purpose and methods used in the study to you after your participation has concluded and to furnish you with our results when they are available. If you have any concerns about your rights or treatment in this or any UBC experiment, you may contact Brent Sauder, Director of the UBC Office of Research Services at You will be given a copy of this consent form for your records.

The manufacturer (Polar™) of the heart rate monitors we use indicates that people with pacemakers or other implanted electromagnetic devices should seek physician's advice before using the monitors. Also, we are interested in testing how the 2Hearts system responds to people with normal heart patterns. **Therefore, if you have a known heart problem, or have a pacemaker or other implanted electromagnetic device, we ask you to please not participate in this study.**

By signing below, I certify that:

- **I have read both (2) pages of this consent form completely**
- **I have received a complete copy of this consent form**
- **I do not have a pacemaker or other implanted electromagnetic device**
- **I do not have any known medical heart problems**
- **I consent to participate in this study under the conditions explained in this form**

Name (PRINT): ______________________________ Date: __________________

Signature: ____________________________________

Project principal investigator:
Sidney Fels
Director: Human Communication Technologies Lab

Appendix E. Copy of Questionnaire Used in 2Hearts--Version 3 User Testing

The questionnaire begins on the next page.

PART I – Listening to / rating audio samples

Instructions:

When you are ready, you will be asked to listen carefully to **2** different musical examples, separated by a short pause. The first of these examples will be referred to as #1 and the second as #2. After you have heard these examples, please fill out your answers to the corresponding questions. Keep in mind that there are no "correct" answers.

The process will be repeated for **4 pairs** of examples (pairs A through D).

Additional Notes:

In some of the questions, you will answer questions using a horizontal-line scale. Please use the scale in the following way:

Place one clear vertical slash mark at any point on the scale, to indicate the direction and relative strength of your answer.

EXAMPLE OF USING THE HORIZONTAL-LINE SCALE:

Q. Which do you like more? Apples or oranges?

I prefer apples.	Neither/the same	I prefer oranges.

|--|--|

Examples_______________

1. Which example would you say sounded "happier"? Make a vertical mark anywhere on the line.)

#1 sounded happier   Neither/the same   #2 sounded happier

|--|--|

2. Which example would you say sounded more "energetic"? (Make a vertical mark anywhere on the line.)

#1 sounded happier   Neither/the same   #2 sounded happier

|--|--|

3. In your own words, how would you say each example made you feel?

#1:
#2:

4. In your own words, describe the differences in sound between the two examples.

#1:
#2:

For each of the 4 phases, indicate whether you agree or disagree with the supplied statement by making one vertical slash mark on each horizontal scale.

Statement: "the sounds drew my attention".

1.

Disagree	**Neutral**	**Agree**

|---|---|

2.

Disagree	**Neutral**	**Agree**

|---|---|

3.

Disagree	**Neutral**	**Agree**

|---|---|

4.

Disagree	**Neutral**	**Agree**

|---|---|

For each of the 4 phases, indicate whether you agree or disagree with the supplied statement by making one vertical slash mark on each horizontal scale.

Statement: "the sounds seemed connected to my mood".

1.

Disagree	**Neutral**	**Agree**

|--|--|

2.

Disagree	**Neutral**	**Agree**

|--|--|

3.

Disagree	**Neutral**	**Agree**

|--|--|

4.

Disagree	**Neutral**	**Agree**

|--|--|

For each of the 4 phases, indicate whether you agree or disagree with the supplied statement by making one vertical slash mark on each horizontal scale.

Statement: “the sounds seemed energetic”.

1.

Disagree **Neutral** **Agree**

|---|---|

2.

Disagree **Neutral** **Agree**

|---|---|

3.

Disagree **Neutral** **Agree**

|---|---|

4.

Disagree **Neutral** **Agree**

|---|---|

Compare and contrast the differences in the **music/audio** produced by the system during each of the 4 phases.

1.
2.
3.
4.

Compare and contrast the differences in the **mood** of your conversation and social interaction during each of the 4 phases. Eg. How were you feeling, how do you think the other person was feeling, how much/how quickly were you talking, etc.

1.
2.
3.
4.

Please indicate whether you agree with the following statement by circling a number.

I liked the sounds produced by the system.

1	2	3	4	5
Strongly disagree	disagree	neutral	agree	strongly agree

Hearing the music distracted me from the conversation taking place.

1	2	3	4	5
Strongly disagree	disagree	neutral	agree	strongly agree

Wearing the heart rate monitor distracted me from the conversation taking place.

1	2	3	4	5
Strongly disagree	disagree	neutral	agree	strongly agree

Overall, I enjoyed using the system.

1	2	3	4	5
Strongly disagree	disagree	neutral	agree	strongly agree

During the experiment I got to know the other subject much better.

1	2	3	4	5
Strongly disagree	disagree	neutral	agree	strongly agree

During the experiment, the sounds produced by the system helped me to get to know the other subject much better.

1	2	3	4	5
Strongly disagree	disagree	neutral	agree	strongly agree

In the space provided, please describe what you liked/didn't like about the sounds produced by the system.

In the space provided, please discuss the following: Would you choose to use this type of system again? What do you like/dislike about the system?

Describe what musical training/experience you have had.

In the space provided, please discuss the following: In what context (place, time, setting) do you think I would be most helpful or enjoyable to use this type of system?

Before participating in the experiment, how well did you know the other subject? (check one box)

☐ I knew them very well

☐ I knew them somewhat well

☐ I met them briefly

☐ I had never met them

Have you consumed caffeine (coffee, tea, or soft drinks) in the previous 24 hours?

Circle: YES NO

If yes, what type of drink(s), and at what time(s)?

____