POLITECNICO DI MILANO
Corso di Laurea in Ingegneria Informatica
Dipartimento di Elettronica e Informazione

MODELING THE EFFECT of SOUND-INDUCED AFFECT in DECISION-MAKING a PRELIMINARY STUDY

Master Thesis
Candidate:
Alessio Contini, ID 754777

Supervisors: Prof. Augusto Sarti
Mrs. Penny Bergman

Academic Year 2010-2011

Contents

Chapter 1

Introduction

The present chapter aims at outlining the role of emotion in cognitive processes such as learning and decision-making,. The latter is a key aspect for the work that has been developed in this study.

1.1 Emotional responses

Emotions are part of our daily life. We frequently talk about emotions, even if it is not trivial to give a unique definition of them. In a review made by Barrett et al. (see [1]) the main streams that have tried to explain what emotions are and how they originate are analized. One approach to the field, called behaviorism, assumes that experiences of emotion are instantiated by physical processes in the brain or body and thus can be completely explained by events in the physical world, i.e. the bodily reaction is the emotion itself. What this theory excludes from the analysis is the actual individual response, lost in the strive for understanding human emotions only in terms of the causes that generate them.

Barrett introduces a different viewpoint: emotions have to be understood both in terms of neurobiological processes and of subjective experience (i.e. asking people what is actually felt).

> ...experiences of emotion are content-rich events that emerge at the level of psychological description, but are instantiated by neurobiological processes, and any theory of emotion experience must address both content and process.
> [1, p. 374]

She proposes a framework in which emotions are seen as a core affect which is enriched by attributes that characterize the specific instances. Such attributes are also imagined as a timeline where events can change the core affect. Core affect and attributes can be associated to particular region activations of the brain, although it is not possible to reduce them to neurobiological activity. In the present study I have adopted Barretts viewpoint and developed a psychological test according to her emotional framework.

The rest of the section will be dedicated to the discussion of interesting aspects in self-assessment of emotional content.

1.1.1 Emotion assessment: dimensions of interest

In order to assess the emotional content of a specific stimulus a method, devised in the framework of a dominant theory called *dimensional theory*, assumes that every evaluation can be brought back to three fundamental measures: *valence*, *activation* and *dominance*. Valence (or *pleasantness*) is the degree by which an emotion is experienced as pleasant or unpleasant along a continuous dimension. According to Barrett valence determines the core affect for the experience of emotion together with activation. Activation indicates the level of excitement or calm associated to the stimulus. Dominance aims at outlining the perception of individual strength with respect to the stimulus, answering the question: does the subject feel dominated by it or dominant over it?

In this work only the two dimensions of Valence and Activation were taken into account, as they represented the two main constituents of the core affect at the basis of emotional experience. From previous research a

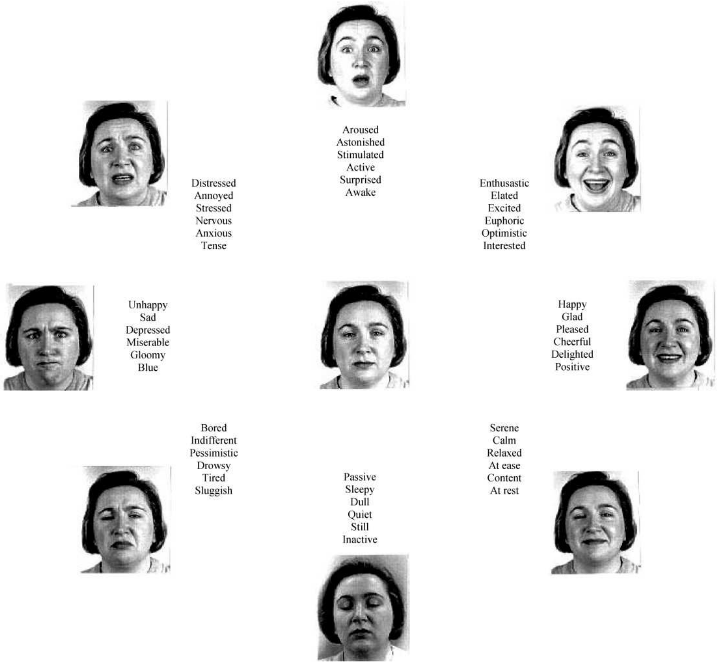


Figure 1.1: The *affective circumplex*.

simple categorization of emotions based on the valence-activation pair has emerged, leading to the results summarized in figure 1.1. This representation can be used as a guideline both for categorizing new sounds and to use old ones in synthesis stages to reobtain a specific emotion.

1.1.2 Emotion and Sound

Evidencies of the fundamental role that the emotional content of auditory events has on us are omnipresent, like in the pleasant favorite song or the scary barking dog. In the acoustic domain the branch that deals with emotional reactions to sound is called *emoacoustics*. While the traditional psychoacoustic studies try to find correlations between perceptual quantities and physical ones, emoacoustic is a branch of it that strives to understand the link between stimuli and the emotional reaction they evoke. Valence and activation are commonly used tools for audio clips rating. Their robustness is such that an American study has been focusing on the creation of a relatively stable database of sound stimuli with their associated valence, activation and dominance mean values and standard deviations. The IADS collection (International Affective Digitized Sounds), is composed by 111 samples of so called ecological sound (natural sounds, not music) ranging from bird singing to women's screams. A selection of auditory clips from IADS has been used in the development of the work (see chapter 3.1).

An interesting aspect that makes room for future, deeper analyses is the fact that ratings of a sound may change depending on the spatial attributes of it and its dynamic behavior (neutral static sources can be perceived as unpleasant ones while approaching the listener, as reported in [2]). The IADS database is composed by static sounds, therefore caution must be put in using them in simulations of moving sources. When a sound does have a emotional content other than neutral, however, it is less likely that its ratings will change substantially.

1.2 Cognitive aspects

Cognitive processes differentiate from emotional reactions in that they are characteristically human (see [3] for a review). Among them the abilities to form a coherent representation of the world, anticipate upon what may be coming next, and make choices about courses of actions play a major role in human activities.

There are four main processes that help us navigating in the reality:

- *interpretation*, that is the ability to extract meaning from ambiguous information;
- *judgment*, consisting in evaluating evidences and giving probabilities to events;

- *decision-making*, the way by which we choose among different options;
- *reasoning*, that is using the information we get for draw inferences.

All of them are complex activities and rely on a series of constituents, ranging from object recognition to activation in semantic memory. Moreover main processes are not to be taken individually, but also connections among them have to be considered. An example is the presence, at the core of decision-making, of a judgement phase in which probabilities of events are estimated.

A very important constituent for all higher cognitive processes is the memory activation. Retention mechanisms allow to elaborate new information based on past experience. A deeper analysis of memory, with focus on the concept of working memory, will be done later in the chapter.

In this framework learning processes, relying on interpretation, judgement and before that on memory, have an impact on decision-making. How learning could be connected to emotion arousal is explained in the next section.

1.3 Emotion and Cognition: the link

Learning activities involve cognitive processes. Then what is the meaning of investigating emotions? An answer came in 1995, when the neuroscientist Antonio Damasio, with his book *Descartes' Error* [4], introduced the hypothesis that emotions produce effects in decision-making and social behavior. Starting from examples of brain-damaged people with lesions in the areas deputed to basic emotion issuing (see section 1.4.1) he noticed they showed impairments in real-life decision-making. In one of his studies published in 1994 together with Bechara [5] it was reported that the patient E.V.R. was:

> ...a prototypical example of this condition. He often decides against his best interest, and is unable to learn from his mistakes. His decisions repeatedly lead to negative consequences.
> [4, p. 7]

This observation led him to substantially contribute to the sociological and psychological research fields. Further evidences were found that areas of the brain which are activated while we are experiencing emotions are the same ones activated when higher level cognitive processes such as decision-making take place. Although emotions cannot be substituted to rational thinking, they represent, according to Damasio, a necessary step toward it.

Blanchette and Richards in their interesting review (see [3]) report evidences that all the main cognitive processes listed in section 1.2 are affected by emotions.

1.4 Learning processes and the Learning stack

The starting point of the present study is the question: *can we find sounds that help people to learn?*

Perhaps trying to find a direct answer goes beyond our capabilities of investigation. What we certainly can do is to trying to chunk down the question in smaller conceptual units and find connections between them. When we talk about *memory*, for instance, we can reasonably assume that its framework of retention mechanisms roots itself in a *learning* activity, where the information are actually captured and processed before being stored. In the same fashion learning is influenced by the substratum of *emotions* (even if many other factors are involved which are out of the scope of the present research).

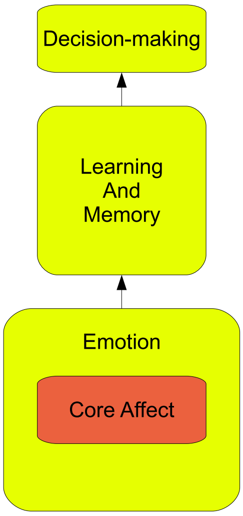


Figure 1.2: The *Learning Stack*.

Following this way of reasoning I came up with a simple graphical schema which illustrates the various levels of what I call *Learning Stack* (depicted in fig. 1.2).

In the next sections each of the elements of the stack will be analyzed with the exception of decision-making, which has been already described together with learning in section 1.2.

1.4.1 Core affect

Core affect can be seen as the seed from where every emotion originates. It consists of a pleasant or unpleasant state that is experienced by an individual. This definition in reported in the thorough analysis performed by Barrett et al. [1], where the main trends in the recent psychological methodology are outlined and the neurobiology of emotions is explored.

When we are in a state of pleasure or displeasure several areas of the brain are activated. The main ones are the temporal lobe with the amygdala (low-lateral parts of the brain), the orbitofrontal cortex or OFC (located just behind the ocular cavities) and the ventromedial prefrontal cortex or VMPFC (situated below the OFC). Together they integrate sensory information, previous knowledge of the experienced object and viscera-motor reactions.

In chapter 2 we will present tools for self-assessment of this pleasant/unpleasant core affect which will be used in the second experiment.

The Pleasantness Problem

Most of the research has focused on the effects of unpleasant stimuli as reported in [6]. One of reasons why this is still the main trend in psychology is the major role from a pathological viewpoint that unpleasantness covers in inducing negative moods, which can account for the creation of many mental disorders. Some attempts have been made to investigate both the polarities and it has been argued that pleasantness is processed through different mechanisms than the better understood unpleasantness (see [7]). In this work I wanted to put the focus on the comparison between “pleasant sounds” and unpleasant ones, which will be protagonists in the experiment, in affecting decision-making (better explained in section 1.2). For doing so the tasks have been chosen as to leave room for different analyses.

1.4.2 Emotions representation

From the core affect the mental representation of emotions relies on three categories of additional content to characterize a specific feeling:

- *arousal* content, connected to the degree of body activation or excitement;
- *relational* content, which expresses the subject’s relationship with another person or environment;
- *situational* content, related to the perception of causal links between a psychological situation and the core affective feeling.

The neurobiology of these attributes is still not understood, even if many studies have been done and some brain areas activation patterns found which pertain to clusters of emotions.

1.4.3 Short-term Memory

A core issue in this work is the understanding of how we process new information and store them for short periods, as we are going to conduct experiments over time-frames of few minutes. As *learning* is the process whose outcome is the *memorization*, we will investigate learning capabilities by assessing changes, if any, in Decision-making games, where past choices knowledge influences future ones. The specific nature of such experiments is thus aimed at exciting those particular mechanisms of our mind which are responsible for the retention and elaboration of information over a short period of time. We can call the ensemble of such mechanisms *short-term memory* (STM).

STM can be subdivided into two main components. The first one is related to the capability of retaining information over time intervals that are less than 1 second for visual stimuli (iconic memory) and around 10 seconds for audio ones (echoic memory). It is called *sensory memory* and cannot be controlled by conscious mechanisms. The second is again called *short-term memory* and has a storing capability inferior with respect to the sensory memory, but can preserve information for time intervals going from few seconds to some minutes, together with processing capabilities. To avoid confusion, from now till the end of the chapter the expression “short-term memory” or “STM” will be used for the latter definition.

In an interesting study by Meiser et alii [8], two main models for STM are compared that have been the main streams along which the cognitive psychology has developed several short-term memory studies: the *working-memory model* and the *changing-state hypothesis*. The focus of the work was the comparison of the effects of secondary tasks on serial short-term memory activities, such as recalling the order by which some visual stimuli were presented.

The original working-memory model, adopted by Baddeley in 1974 (see [9]), distinguishes a *central executive* memory system from a set of modality-specific subsystems, each one of them dealing with a particular category of stimuli. The central executive system (CES) acts as a processing unit of mixed information. The two main subsystems, the *phonological loop* (presumably related to language learning) and the *visuospatial sketchpad*, provide the CES with the data to be processed.

Every system has limited capacity, thus overloads in a system, either central or peripheral, can occur. If the stimuli fall under one category, only the related modality-specific subsystem is saturated, causing no detrimental effects on other subsystems. Therefore no significant cross-modality interferences are expected. A successive extension of the model (see [10]), adds a fourth component, called *episodic buffer*, to account for cross-modal stimuli integration.

The changing-state hypothesis states that cross-modality effects can occur, as a unitary short-term storage is hypothesized. This implies, for example, that secondary task involving auditory stimuli would load the common storage together with primary visual task, leading to a potential saturation of the STM and a consequent decrease in performance.

1.5 Hypotheses

As we have seen so far emotions affect cognitive processes, including memory, learning and decision making. We further have evidences that sound can trigger an affective response (see [11] [12]). Auditory stimuli should thus have the ability to affect cognitive processing. The question arises of whether or not we can find evidences supporting this conjecture. In order to investigate the answer to this question we have to look more closely at what behavioral aspects are modulated by affect.

According to Blanchette [3] while we are in a positive mood we tend to increase the risk aversion in decision-making. Hence, the hypothesis is:

> ***Hypothesis 1:***
> ***Positive mood-inducing sounds increase risk aversion in decision-making.***

On the other hand negative mood can lead to an increase in risk taking. Thus my second hypothesis is:

> ***Hypothesis 2:***
> ***Negative mood-inducing sounds increase risk taking in decision-making.***

Testing these hypotheses could give us more insight into the issue. What the next chapter will deal with are the tools and methods of investigation used for addressing the problem.

Chapter 2

Methods

2.1 Psychometrics

2.1.1 Games as abstraction of reality

We can find in every culture certain games that are not just meant for amusing or entertaining, but also as teaching tools to foster the learning activities in a light and yet effective way. Their strength relies in the fact that they start from a simplified vision of reality and build on top of it tasks to cast light only on specific aspects. Under this perspective they offer us a *powerful abstraction of the real domain.*

In the scientific environment games arose such an interest that in the XXth century even a set of formal methods to investigate their structure, possible solutions and mechanisms, called *Game Theory*, started developing. Nowadays Game Theory has penetrated into many fields, ranging from social sciences to economic models[1]. One of the constituting axioms, called of *rationality*, states loosely speaking that individuals always follow a strategy from which they can have back the maximum personal utility and this concept, according to Blanchette and Richards [3], reflects, to some extent, in real situations. Game theory classic formal structure is mainly focused on games with several participants.

On the other hand in the neuro-behavioral tradition several tasks have been developed with a tight connection to the mechanisms that instantiate emotional and cognitive reactions at the brain level. This aspect persuaded us to embrace a more psychological-based view of games for the present study and use, as a starting point for the design and implementation of the test, what perhaps is the most famous one since 1995 is Bechara's *Iowa Gambling Task* (IGT), described in the next section.

---

[1] for an introduction to Game Theory (in Italian), see [13]

2.1.2 The Iowa Gambling Task (IGT)

If we want to catch some details of our Decision-making strategies we can devise games whose outcome depends on user choices. It is also crucial to keep the playing mechanism simple and intuitive both for having more chances of getting unbiased data and for easing the process of hypotheses elicitation. Card games are a perfect candidate for addressing these core issues and among them gambling tasks give lot of flexibility in rule definition and level of engagement. Under this category the Iowa Gambling Task (IGT) falls.

The IGT was designed in order to mimic the gains and losses coming from decisions made in real life situations. Indeed every choice bears with itself both positive and negative aspects and the goal of this card game was to account for this two-faced nature of decision-making.

In the original Bechara and Damasio's study [5] four decks of cards where presented side by side and facedown. Each deck was composed by winning and losing cards in different proportions to have two decks that can be regarded, on average, as winning and two as lossy ones. The subject was given an initial credit to play with and asked to draw a card from one of the decks. Once the decision was made the result was displayed and the credit updated accordingly. The game then prosecuted in the same fashion until a certain number of cards (typically 100 or 150), were drawn. There was no limit on the number of cards that could have been selected from a single deck.

Although the task was originally designed to test differences between clinically normal people and brain damaged subjects, a whole stream of research was born around it for investigating several aspects of inter-group differences (drug abusers vs. normal people [14], ethnic peculiarities [15], task performance related to people with different perception of wins/losses entities [16]).

2.1.3 The version of the IGT in the present study

The design of the task for the test (see chapter 3), was based on a modified version of the IGT proposed by Peters and Slovic [16]. In their research they have introduced some interesting concepts that improve the original task:

- the card decks are displayed one at the time, forcing the subject to evaluate every option and lowering the risk of getting stuck from the beginning in choosing a single deck instead of exploring them all;

- when a deck is presented the user can either draw a card from it or switch deck and in the second case the card is not discarded: the sequence of positive and negative results is kept consistent with the expected average outcome of each deck.

Such outcomes were pseudorandomly scrambled to ensure the right proportion of wins and losses. Decks were grouped into three different categories in order to enrich the spectrum of possible analysis:

- winning or “good” decks and losing or “’bad’ decks, where the average results were postive and negative, respectively;
- high wins and low wins decks, where the magnitude of wins was substantially different;
- high losses and low losses decks, where losses had different magnitudes.

	Decks			
	A	**B**	**C**	**D**
Good	No	No	Yes	Yes
High-loss	No	Yes	Yes	No
High-gain	No	Yes	No	Yes

Table 2.1: Deck grouping in the second game

The way the four decks were grouped is summed up in table 2.1.

2.1.4 Unipolar and Bipolar measures

Among self-reported psychological dimensions one of the most widespread techniques is the rating scale: an oriented line along which the subject checks the point that better represents his emotional state. The two typologies of scales we can find are the unipolar and bipolar ones. We are using a unipolar scale when a single concept representing a dimension is provided (e.g. being satisfied) and the edges of the line are the two extremes of the intensity perception (not at all/completely). In a bipolar scale two contrasting concepts are put at the edges (e.g. excited/calm), and each point on the scale represents a mix in different proportions of the two.

Valence-Activation ratings: SAM scale

In this work we use self ratings principally for assessing the perceived valence and activation of the soundscapes, both before and after the cognitive task. Various attempts have been made to find suitable tools for evaluating this orthogonal dimensions. In the ’70s a popular system was the Semantic Differential scale (SD scale), devised by Mehrabian and Russell, which required 18 different ratings. Although the method was informative, it represented, as stated by Bradley and Lang,

> [. . .]a heavy investment of time and effort, and results in a relatively large database that requires statistical expertise for

> resolution (i.e. factor analysis) [...] reliance on a verbal rating system makes it difficult to utilize this methodology in non-English speaking cultures (unless there has been translation and validation) and in populations which are not linguistically sophisticated (e.g., children, aphasics, etc.).
> [17, p. 50]

A more parsimonious way of looking at this 3-dimensional structure was provided by Lang, who developed a graphic-based assessment tool to directly rate the interesting indexes; he called it *Self-Assessment Manikin* or SAM. It is made of three graduated scales with illustrations to recall the different dimension to be evaluated. Thanks to the graphical representation the language issues were overcome, while keeping a good correlation with the measures obtained through the SD scale.

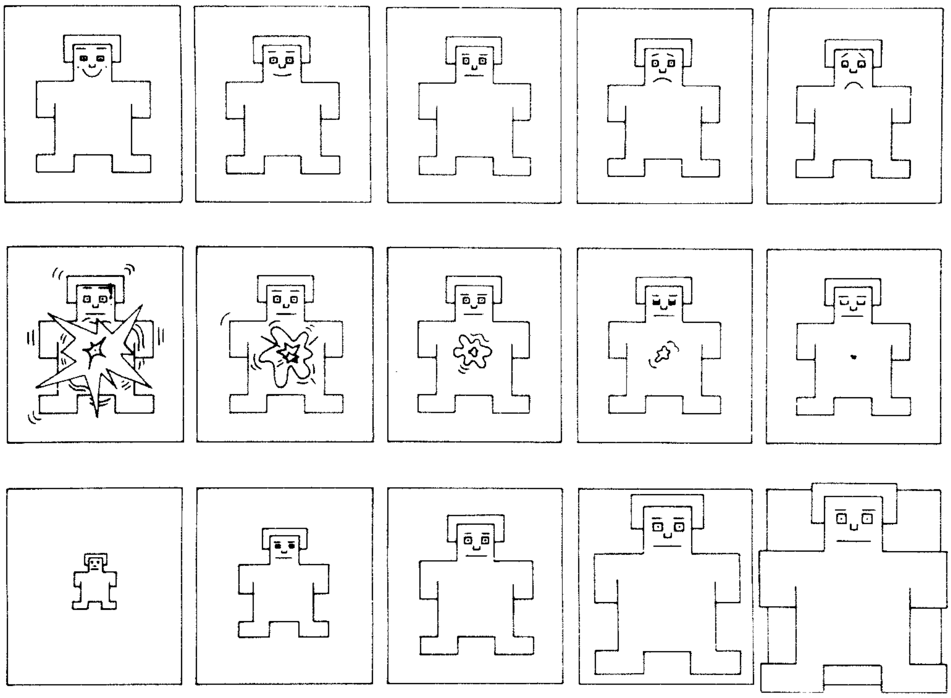


Figure 2.1: The SAM scale illustrations for the three dimensions.

The SAM scale illustrations in a Lang's 1985 article [18] is shown in figure 2.1. From left to right the first strip represents the valence dimension, from pleasantness to unpleasantness. Similarly the second strip pictures activation, from excitement to quiet. The last strip shows the dominance dimension, from dominated to dominating.

The measures of dominance on the SAM scale had the least correlation with the SD scale as stated in [17]. This is another reason why only valence and activation were assessed in the present study.

Demand, Scare and Familiarity ratings

Together with the two SAM measures other ratings were used with the purpose of better characterizing the experience of emotion of each subject, both for themselves (being consequently more precise in devising the perceived valence and activation), and for us. While the SAM ratings are bipolar, the other three grading scales were unipolar, ranging from [not at all] to [a lot]. The first dimension was the level of *demand*. It grades how much the stimulus is attention catching for the subject. The experiments were designed in such a way (see chapter 3), that the gambling tasks had the mood-inducing sounds playing in the background. The effect of this task-irrelevant audio could be evaluated in terms of demand. In fact several studies have hypothesized a disruptive influence of irrelevant sound on serial recall (see [19] as a recent example). In our card games we do not look for conscious serial recall, but the order of appearance of cards and their associated values are crucial to subconsciously learn the difference between the decks. Moreover a detrimental effect in game performance could support the Changing State Hypothesis as exposed in chapter 1 against the Working Memory model.

The level of *scare* assesses how fearful a stimulus is. This indicator has been introduced to stimulate the subject to give a more complete description of the emotional content of the sound.

In the same fashion the measure of *familiarity* of the stimulus investigates how much the subject is used to that sound. It is reasonable to expect a change in familiarity as the task proceeds.

2.1.5 Discarded measures

Besides the game itself, together with data regarding the choice history of every participant, some other measures are typically collected, because they enrich the spectrum of possibilities for the hypotheses to be stated and for the data analysis stage. Some candidates are physiological(skin conductance, mio-sensors for muscle contraction), other self-seported (questionnaires). The latest have been used extensively in psycho-sociological studies and many methods have been developed for their analysis. Even if the instrumentation for skin conductance measurements was present at the department, the choice of using touchscreens as interaction devices prevented me from inserting those sensors. In fact:

- they could have been placed on fingers, either of hands or feet;
- the hands must have been left free to use the touchscreen;
- the tests were run in winter and exposing feet to the cold environment would have been a strong source of bias in mood-induction;

- in a previous work at the department showed that those measurements were affected by noise, to the point they had not been used in data analyses;
- coupling decision-making patterns with self-reported measures would have given us a more complete vision of the problem, both from behavioral and emotive standpoints as suggested by Barrett et al. in [1].

2.2 Soundscapes: binaural synthesis

The challenge of mood induction through acoustic stimuli presents several aspects. The first to be taken into account is the fact that a single typology of audio samples is hardly capable of a positive mood induction. Even the most pleasant sound event may be annoying when put in a loop and heard hundreds of times consecutively. In real life we are typically immersed in a continuously changing auditory environment. Under this perspective, when trying to engage people with complex sound stimuli, one of the approaches which makes use of "ecological" sounds (those ones we hear in our environment naturally, from a car motor to a bird singing), is the one of *soundscapes creation.*

We can think of a soundscape as the acoustic counterpart of a landscape: several auditory images emerge from a background environment, giving us a sensation of immersiveness. The more the synthesized scenario is familiar to the listener, the easier it is to convey an emotional content. If this effect is achieved, inducing an emotional reaction in the listener becomes easier than trying to obtain the same effect with single audio stimuli.

2.2.1 The third Hypothesis

The goal which was set due to these considerations was to reproduce a plausible sound environment with a certain underlying pleasant/unpleasant impact on the subject. The assumption we made was that using in the creative process some key sounds with known valence and activation would have influence the overall emotional connotation of the auditory scene. We therefore stated a new hypothesis:

> ***Hypothesis 3:***
> ***The emotional content of a soundscape can be predicted by averaging the emotional content of sounds previously categorized according to specific affective ratings mechanisms.***

2.2.2 Rated sounds and synthesis method

Such sounds were selected among the ones of the IADS database. As the semantic content of such sounds was quite heterogeneous we opted for creating a sort of auditory story in which the subject is the protagonist and moves from one key sound to another through smooth synthesized-environments transitions.

In order to render with a good approximation the spatialization of sounds while keeping the setting of the experiment simple I opted for *binaural synthesis*. This approach has the advantage of requiring a simple headset for the reproduction of the sounds, obtaining a good 3D audio effect with a stereo track. In the following subsections the main principles of binaural synthesis will be mentioned and an overview of the mathematics behind the synthesis method I adopted is provided.

2.2.3 Sound localization and HRIRs

Sound localization is one of those topics that have been investigated for a century. In human beings it is affected by many physical factors, above all the shape of the outer ear. It also depends on reflection patterns on shoulders and torso, together with the shape of the head. The *pinna* (the outer ear cartilage), has a fairly complex shape which alters the incoming sounds depending on the direction of arrival and on the frequency. The distance of an acoustic stimulus from the receiver is another factor that can be estimated by reflection patterns perception or direct evaluation. We also have to consider that pressure waves decrease in amplitude due to air absorption with the inverse of the distance between source and listener.

Under an engineering perspective we can model all these effects through a unique linear time invariant filter. This representation allows us to apply classic Fourier theory of signal filtering and interpret the sound that reaches the ear canal as the convolution between the sound at the source and the linear filter that models the impulse response of the transmission channel from source to receiver. We call this impulse response Head Related Impulse Response (HRIR) and its spectral-domain equivalent Head Related Transfer Function (HRTF). It has to be noticed that HRIRs come in pairs (for left and right ears). Fixed the listener, for each position of the source in the three-dimensional space we have a different pair of HRIRs.

It is possible, given the HRIRs of a specific position in space and an anechoic sound, to synthesize the spatialized sound. The next section will deal with two approaches for obtaining the HRIRs.

2.2.4 Synthesis approaches

HRIRs can be either measured from real subjects (or mannequins), or be synthesized through a model of the structure and interactions of pinna, torso, head.

With the first approach we have to sample the 3D space in several locations and for each one measure the impulse response by emitting a signal and capturing it with microphones which are usually placed inside the ear canal, so the sound will be recorded once the filtering effect of the pinna has already taken place. This kind of measures are conducted in anechoic environments, where no reflection paths should occur. As we record only a finite number of HRIRs and we do that for a fixed distance, we will have to interpolate the impulse responses in order to approximate them in position other than the sampled ones.

The second approach tries to find suitable general or adjustable-parameters methods for creating synthesized HRIRs. We can obtain fairly good spatialization effects by only playing with auditory stimuli loudness (a sound of increasing loudness is perceived as approaching), and phase difference in headphones (delaying one of the two channels causes the sound to be perceived outside the head in a direction on the horizontal plane depending on the entity of the delay). Such expedients fail to provide the listener with cues for discriminating between the frontal and rear halfspaces and cannot provide any elevation information.

In the 1920s the basic relationships between pure tones perceived position and intensity and phase differences were already known (see Firestone [20]). Researchers were also aware that modeling the head as a sphere led to estimation errors of increasing magnitude as the frequency of the tone was raised. Firestone found that such errors were negligible within $1KHz$, but already at $2KHz$ the difference between predicted and measured phase differences and amplitude ratios of the sounds became substantial. To improve the quality of sound localization several details must be taken into account, starting from a different modelization of the head, the influence of torso and, above all, the outer ear structure.

2.2.5 Model used in the present study with sampled HRIRs

For synthesizing the soundscapes it was decided to use already sampled HRIRs of a human individual. Suitable sets were found in a database made by Ircam in 2003, where 51 individuals' impulse responses from 187 position in 3D space were acquired (see [21]). The website also allows to download sample Matlab code for rudimental synthesis. Although the shape of the pinna changes from person to person and makes impossible to achieve the same subjective auditory realism spatializing the sounds with someone else's impulse responses, a fairly good effect can still be achieved. Therefore every set of measurements was equivalent to another under the realism aspect and I chose a set of them labeled with id 1047 from the aforementioned collection.

Now we will examine a model for binaural sound synthesis based on convolution by sampled HRIRs when both source and listener are fixed and derive a simple method to deal with moving sources and fixed head. No Doppler effect is considered in the moving source model.
Let $s_{AN}(t)$ be a continuous-time sound signal recorded in an anechoic environment and emitted from an ideal static point source. Let $h_R(\overline{v}, t)$ and $h_L(\overline{v}, t)$ be the HRIRs of right and left ear respectively, where $\overline{v} = \overline{x} \cdot d$, with $\overline{x}$ the *position versor* that from the subject points toward the source and d the *euclidean distance* between the source from the center of the head. An expression for the binaurally synthesized sounds $y_R(t)$ and $y_L(t)$ would be

$$y_R(t) = s_{AN}(t) * h_R(\overline{v}, t), \quad y_L(t) = s_{AN}(t) * h_L(\overline{v}, t). \tag{2.2.1}$$

Moving to the frequency domain

$$Y_R(f) = S_{AN}(f) \cdot H_R(\overline{v}, f), \quad Y_L(f) = S_{AN}(f) \cdot H_L(\overline{v}, f). \tag{2.2.2}$$

Each term $H_i(\overline{v}, f)$ is rewritten as

$$H_i(\overline{v}, f) = H_{ff}(d, f) \cdot W_i(\overline{x}, f), \tag{2.2.3}$$

where $H_{ff}(d, f)$ incorporates air damping and propagation delay and can be called the *free-field transfer function*; $W_i(\overline{x}, f)$ is the component of the HRIR responsible to give the perceived direction of arrival of the sound by filtering the incoming signal, i.e. the *ear transfer function.* We have then factorized the HRTFs into a part which is common to the two ears (the term only dipending on distance), and the two related to the close-to-the-ear filtering.

We can then write the two outputs as:

$$Y_i(f) = H_{ff}(d, f) \cdot [S_{AN}(f) \cdot W_i(\overline{x}, f)] = H_{ff}(d, f) \cdot V_i(\overline{x}, f), \tag{2.2.4}$$

where $V_i(\overline{x}, f)$ represents the close-to-the-ear filtered sound, a *normalized monaural synthesis.* When using sampled HRIR we only have a reference distance between the test source and the receiver.

Assuming a constant phase delay of air filtering (no frequency dispersion), and a decrease of sound pressure with the inverse of the distance source-listener, $H_{ff}(d, f)$ can be written as:

$$H_{ff}(d, f) = \frac{1}{d} \cdot e^{-jkd} \text{ , with } k = \frac{2\pi f}{c} \text{ ,} \tag{2.2.5}$$

where c is the sound propagation in air. Thus the free-field transfer function is a block that changes the perceived distance of the source with respect to the reference one (in Ircam database all the measurements were made at

1 meter of distance from the listener) by varying the amplitude according thanks to a term $A(d) = 1/d$ and introducing a frequency-independent time delay $\tau = d/c$. Rewriting equation (2.2.5) in terms of $A(d)$ and τ

$$H_{ff}(d, f) = A(d) \cdot e^{-j2\pi f\tau} \tag{2.2.6}$$

and going back to time domain

$$\begin{aligned} y_i(t) &= h_{ff}(d,t) * [s_{AN}(t) * w_i(\overline{x}, t)] \\ &= [A(d) \cdot \delta(t-\tau)] * v_i(\overline{x}, t) \\ &= A(d) \cdot v_i(\overline{x}, t-\tau) \end{aligned} \tag{2.2.7}$$

According to equation (2.2.7) we can first convolve the original signal by the close-to-the-ear impulse response and then apply a scaling by $A(d)$ and a time delay of τ. But scaling and time delay are operations that can be easily performed with simple audio editing tools such as Audacity. We can then:

- perform the convolution operations with softwares as Matlab (this has been the case of the present study, where the HRIRs where provided as Matlab data);
- simulate distance cues through manipulation of the convolved sound with audio editing tools (I have used Audacity).

Although tuning manually amplitude and delay does not seem to be useful in static source conditions, things change substantially when we introduce a dependence on time for distance and direction of arrival (i.e. we have $\overline{x} = \overline{x}(t)$ and $d = d(t)$). If we keep divided convolution and distance scaling we can first create a normalized binaurally-synthesized moving source by interpolating the HRIRs of the directions we span with the trajectory and then move to the audio editing tool for the insertion of the time delay and the amplitude modulation for mimicking the source distance. If this approach is pursued we introduce an approximation with respect to the time delay: we assume it to be constant. Under the assumption of a slowly moving sound source ($c_{source} \ll c$), though, a constant delay does not alter the perceived sound movement substantially.

This model has been used in the synthesis of soundscapes.

2.3 Statistical tools

The basic statistical tools used in the study are hereby listed with a brief introduction.

2.3.1 t-test

In order to evaluate how the mean values of two groups relate to each other a suitable metric is the computed value of a *Student's t variable*, obtained by dividing the difference between the two sample means by an average sample *standard deviation* $\overline{s}$ (the square root of the variance). In formula:

$$t = \frac{\hat{\mu}_1 - \hat{\mu}_2}{\overline{s}} \,. \tag{2.3.1}$$

We can interpret this probability as the *likelihood that two sample means are drawn from the same sampling distribution* and any difference between the drawn samples is merely casual. The underlying assumption of the test is the normality of all the variables involved. A threshold likelihood of 5%, meaning that only 5% of the values of the t distribution can be regarded as been drawn from the same normal distribution, is commonly adopted to discard the hypothesis of equality of means and to state that their values differ significantly. Another value of interest, stricter than the previous one, is 1%.

The computed likelihood is called *p-value*, while the method of the comparison of two means is known as t-test.

2.3.2 ANOVA

Despite its name, the Analysis of Variance method (ANOVA) for significance tests is based on comparisons of mean values of variables with respect to some parameters. Thus the variance can be thought of as the quantity to be investigated in order to extract information about how the means of different groupings are related to each other.

We can regard the ANOVA method as a generalization of the t-test to more than two groups. ANOVA evaluates the likelihood that the means of *two or more groups* belong to the same sampling distribution. For doing so it relies on a subdivision of the total *sum of squares* SS_{TOT} (the variance multiplied by the number of samples minus one), into partial sum of squares accounting for the effects of the parameters. It is in fact possible to separate the SS_{TOT} in two main terms:

- a sum of squares SS_B accounting for *between-group variability*;
- a sum of squares SS_W accounting for *within-group variability.*

We would expect, if the values of the samples in one group are well clustered around their mean, to have a small value of SS_W compared to SS_B, supposed to represent the degree to which the means of the groups are separated from each other.

It is therefore natural to evaluate the ratio of the two and speculate on its value. We use the following metric:

$$F = \frac{SS_B}{SS_W} \ . \tag{2.3.2}$$

As in this case we evaluate the ratio of two variations instead of the difference of two means, another statistical test, called F-test, is used. As we did before, we define a null hypothesis H_0, stating that *all the means of the different groups are equal.* A small likelihood of H_0 implies that *at least two of the means are significantly different.* A α-value of 5% is again a commonly accepted significance threshold for the test.

ANOVA is one of concepts at the basis of the *General linear models* analysis that has been performed in the study and discussed later.

2.3.3 General linear models

General linear models incorporate many statistical analysis tools such as ANOVA, MANOVA, ANCOVA, MANCOVA [2] , t-test. Their mathematical definition is an extension of *multiple linear regression.* Recalling that a multiple regression problem consists in finding the coefficients β_i that relate the dependent variable y to the independent variables x_i in the following equation:

$$y = \beta_0 + \beta_1 \cdot x_1 + \ldots + \beta_I \cdot x_I \ , \tag{2.3.3}$$

we can solve the problem of estimating the coefficient β_i of the equation by collecting several input vectors $X_j^T = [1, x_{j1}, x_{j2}, \ldots, x_{jI}]$ and their related outputs y_j. We put

$$Y = Xb + e \ ,$$

where

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_J \end{bmatrix} , \ b = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_I \end{bmatrix} , \ X = \begin{bmatrix} X_1^T \\ X_2^T \\ \vdots \\ X_J^T \end{bmatrix} , \ e = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_J \end{bmatrix}$$

The solution of this system is readily find if we assume $J > I$ and matrix X full rank (i.e. the columns are independent), and it is found through the *least square estimator*:

$$\hat{b} = \left(X^T X\right)^{-1} X^T Y \ . \tag{2.3.4}$$

Multiple regression has two main issues:

- it is limited to a single dependent variable analysis;

---

[2] MANOVA: Multivariate ANalysis Of VAriance; ANCOVA: ANalysis of COVAriance; MANCOVA: Multivariate ANalysis of COVAriance

- it does not provide a solution when X is not full rank.

This problems are overcome by general linear models. The vector Y can be extended and become a matrix whose columns represent different dependent variables. Moreover the concept of *generalized inverse* of a matrix is introduced: if the matrix is full rank the generalized inverse reduces to the regular inverse. In general linear models we can also make analysis taking into account categorical data such as gender: they are converted in numbers for the sake of calculations. The general linear model can be expressed in equation form as:

$$YM = Xb + e \;, \tag{2.3.5}$$

where Y is the matrix of dependent variables, M is a trasformation matrix which determines linear combinations of the dependent variables, X, b and e are as described for the multiple regression model. The solution of the system is given by

$$\hat{b} = \left(X^T X\right)^{-1} X^T Y M \;. \tag{2.3.6}$$

In equation 2.3.6 the matrix $\left(X^T X\right)^{-1}$ is the generalized inverse if X is not full rank.

Within-Subject (Repeated Measures) Designs

One of the interesting aspects of general linear models is that they allow us to analyze data consisting in repeated measures of the same dependent variable under different control conditions. In this work the test compared the choice patterns of an individual exposed to two different mood-inducing soun stimuli: this situation is an example of repeated measures design. General linear models have been the main statistical tool in the present research.

2.3.4 Effect size: partial eta-squared $\eta_P{}^2$

In psychoacoustic tests with several parameters the significance levels do not exhaustively show the interaction effects, as we do not know in which proportion a single parameter influences the variability of the samples. It could happen that the majority of the differences are due to noise and only a small percentage is related to a real effect: this would cast a shadow on the analysis performed, showing that a significant result could not have a noticeable impact on a global scale.

In order to account for this potential issue new metrics for assessing the *effect size* are introduced. In particular we will use a quantity, called *partial eta-squared* or $\eta_P{}^2$, that is defined as:

$$\eta_P{}^2 = \frac{SS_{effect}}{SS_{effect} + SS_{error}} \;,$$

where:

- SS_{effect} is the variability attributable to a parameter;
- SS_{error} is the error variability attributable to the same parameter.

This index is by definition a proportion of the single effect size over the error size. The interpretation of η_P^2 is related to the effects we are probing. As written in [22]:

> A small effect may have important consequences if it distinguishes two models, while a large effect may not matter if it was already expected under any theory.

It must be noticed that this index is not the only one developed in order to address the issue of estimating the effect size of a variable. Other metrics, such as η^2 (*eta square*, not partial eta square) and Cohen's d are used in different situations (for a review see [22]).

In recent years a dispute in the statistical world has seen two different standpoints arguing on which approach should be preferred when evaluating the outcomes of an experiment and drawing conclusions from the analysis of them. Some scientists regard the significance level of a test as the main parameter to judge its quality, while others think that a relatively high effect size could still be indicative of an important correlation between variables even when significance levels would go against its existence. In the present study we decided to take into account both of them in the statistical analysis.

2.3.5 Validity and Reliability

The concepts of reliability and validity are extremely important in what concerns the statistical analysis of data. Especially in psychological studies one of the most challenging issues is to devise psychometrics and tools whose outcomes remain stable if the tests are repeated several times under the same conditions. Reliability is the consistence of the results when repeating the assessments. We talk about validity when, if we have a way of grading the outcomes of a test, such measurements state the right value. In the present study we use decided to rely on bipolar scales (especially valence and activation ones), and on a modified version of the IGT which have been extensively tested in order to assess both reliability and validity (as in [16] [23] [12] [24]).

Chapter 3

Test implementation and execution

3.1 Soundscapes realization

For creating the soundscapes I started from a selection of IADS samples, imagined a story with few main events and tried to render it binaurally. Each of the two stereo tracks consisted of 120 seconds in which several auditory images were concatenated together in an organic story. A longer version of each soundscape was obtained by concatenating it three times and slightly modifying the key sound disposition.

3.1.1 Procedure and tools developed

Ircam functions modification

For realizing the soundscapes I developed a set of functions for synthesizing binaural signals. The original Matlab functions provided with HRIRs by Ircam (in particular the one in charge of performing the binaural synthesis), were a good starting point for understanding the procedure and adapt it to my needs. The available function for three-dimensional synthesis of sounds took as arguments the azimut and elevation of one point on the sampled sphere, together with a mono track sound sample ad the specific individual's HRIRs for every sampled position. The output is a couple of mono signals, one per ear, which were the original signal convolved by the selected pair of impulses responses.

The module however was quite rudimental and it did not allow to pass as an input a series of points of a hypothetical trajectory in space to be simulated. For overcoming this difficulty I rethought the synthesis mechanism adding to the input parameters the name of a text file in which, for every row, the location of the sampled point and the time (in sample), of the beginning of the convolution are stored. The new function chunked down the original signal in blocks that were convolved with different HRIR pair depending on the position in the trajectory. Sometimes audible artifacts were generated in the transition from one position to another; this problem was overcome during the editing of the soundscape through Audacity.

Sound provenience

The second step has been the sound retrieval; three sources for the sounds were used:

- the key sounds from the IADS database;
- self-recorded sounds;
- sounds downloaded from an on-line open database.

The IADS were the fundamental audio clips that have been "glued together" by means of other sounds. All the self recorded sounds were acquired in WAV format with a Zoom H2 microphone at a sampling rate of $44100Hz$. Sounds from on-line databases were selected according to the quality of the recording. Although almost none of the sounds, regardless of the provenience, had not been acquired in an anechoic room, the environmental bias introduced was negligible (most of the time the recordings had been made in open spaces, where reflections are almost non existent).

3.1.2 Pleasant sound creation

The IADS samples for the pleasant soundscape are listed in table 3.1.

IADS samples - Pleasant soundscape

Sound	**IADS ID**	μ **Val**	σ **Val**	μ **Act**	σ **Act**
Baby laugh	110	7,92	1,55	6,04	2,08
Cardinal	151	7,35	1,64	2,73	2,00
Applause	351	7,69	1,40	5,77	2,26
Roller coaster	360	6,90	2,15	7,36	1,90
Beer	721	7,13	1,45	4,43	2,16

Table 3.1: IADS samples selected for the pleasant soundscape

They all have valence values around 7 and different levels of activation. Moreover they are quite different from one another semantically speaking. I grouped them in three different contexts: an amusement park, a quiet wood and a cosy pub. Under this perspective I then imagined a story to link all the sounds and the contexts together. The final sequence started from an amusement park with roller coasters, people clapping to clown performances and a baby laughing, moved to a wood with Cardinal birds singing and ended up in the pub with a bottle of beer.

According to the "screenplay" I retrieved some accessorial sounds to be used as either background or link between the different moments. The selection is proposed in table 3.2, where it can be noticed that only one sound among the ones used was self-recorded. I deliberately inserted again sounds that were already among the key ones. This was done because the IADS samples were too short to be fruitfully used as they were in an attempt to make the auditory scene plausible; in particular the roller coaster in the IADS was few seconds long, in contrast with what really happens in an amusement park. The new sounds have been used as a semantical reinforcement for the IADS ones.

Downloaded sounds	**Recorded sounds**
Bicycle passing by	Walk in a wood
Pub environment	
Applause	
Trumpet in circus	
Roller Coaster	
Walking in a park	
Door shut	

Table 3.2: Recorded and downloaded samples for the pleasant soundscape

3.1.3 Unpleasant sound creation

An identical procedure was used for the unpleasant soundscape elaboration. The IADS sounds are listed in table 3.3. They all have a quite low level of valence except from the helicopter that has been inserted to compensate for the screaming woman clip, which had been rated as the most unpleasant of all. The plot in this case includes two main context: the walk in the park and the chaotic city. Starting from a quiet walk that turns into a nightmare because of a swarm of bees the scene moved toward the city and the daily road traffic with an aggressive dog growling and a baby crying while passing close to a construction site. After having crossed the road an helicopter

IADS samples - Unpleasant soundscape

Sound	**IADS ID**	**μ Val**	**σ Val**	**μ Act**	**σ Act**
Dog growl	106	2,73	1,92	7,77	1,71
Bees	115	2,58	1,45	7,01	1,93
Baby cry	261	2,84	1,61	6,49	1,66
Fem Scream	277	1,74	1,16	8,07	1,40
Helicopter	410	5,19	1,79	5,50	1,96

Table 3.3: IADS samples selected for the unpleasant soundscape

passes above the fictitious protagonist and the scenario ends with a woman screaming in an alley.

This time I only used self-recorded sounds in the soundscape synthesis, together with the IADS samples, as shown in table 3.4. Road traffic was used as background sound for most of the scenario and few elements in it were convolved with HRIRs to give the impression of cars approaching the listener.

Recorded sounds
Walk on dead leaves
Run on dead leaves
Pile drivers
Tram stopping
Traffic noise

Table 3.4: Recorded and downloaded samples for the unpleasant soundscape

3.2 Game realization

Some of the main features of the version of the IGT I have implemented were discussed in section 2.1.3. Here we go more into details explaining the game dynamics and the main choices performed in terms of technologies and realization.

3.2.1 Game structure

The game, based on the work of Peters and Slovich (see [16]), consisted in a succession of trials. In each trial one of four decks of cards was presented and the subject was prompted to choose whether to draw a card (they were displayed facedown), or to switch deck. Switching did not imply discarding the card from the deck, therefore the next time the same deck would have appeared a new choice would have made on the same card.

The four decks were composed by mixes of winning and loosing cards. Magnitude of wins and losses, probability of drawing a winning card in a specific deck and the expected outcome of a card are reported in table 3.5. A detailed list of characteristics for each deck is presented in table 3.5.

It is important to point out the difference existing between display of decks and order of the cards within each deck. In this task this two orthogonal concepts had to be taken into account. Every deck should have been displayed with the same frequency in the game and with a fairly even distribution along trials. In the same fashion also the frequencies of wins and losses needed to be quite regular for avoiding extreme situations, such as the occurrence of mainly good cards at the beginning of the game and

Payoff Variables	**Decks for the test**			
	A	**B**	**C**	**D**
Gain Amount	50-150	150-250	50-150	150-250
Average Gain Amount	100	200	100	200
Loss Amount	100-200	200-300	200-300	100-200
Average Loss Amount	150	250	250	150
Probability (Gain)	0,5	0,5	0,8	0,5
Expected card value	-25	-25	30	25
Coding of good decks	0	0	1	1
Coding of high-loss decks	0	1	1	0
Coding of high-gain decks	0	1	0	1

Table 3.5: Deck grouping in the second game

of negative ones when approaching the end: the bias introduced would not have been negligible.

In the game the display of decks occurred pseudo-randomly, in that the trials were divided in groups of 8 displayed cards, 2 for each deck. Within each group cards were scrambled so to have a slightly different order of display. It this way we had a good balance between frequency of deck appearance and randomness in trials. Concerning the consistency in proportion of wins and losses within decks, every 10 cards the average win corresponded to the average win of the deck and the proportions of winning and losing cards were adherent to the probabilities in table 3.5. Each game came to an end once 100 cards were drawn, all of them potentially from a single deck as when rejecting a card the count of drawn cards did not increment. Each deck, then, was composed by 100 cards.

In order to avoid deck-related bias at the beginning of the game, the first 2 cards of each deck always led to a win.

3.2.2 Game interface

Touchscreens as interaction devices

The implementation of the IGT was computer based and developed within a software environment for neurobehavioral tests called Presentation. Among several options in how to design the interaction of the user with the application I decided to substitute the classical ones such as keyboard and mouse with an approach based on the use of touchscreens. Apart from personal curiosity, the advantage with touchscreens is that they can better mimic a real game where I actually select by hand a card. Moreover using touchscreens and not, for instance, keyboards, allows the experimenter to reduce the occurrence of accidental interaction of the user with the system; such events could happen with Presentation, where pressing the escape button causes the experiment to interrupt.

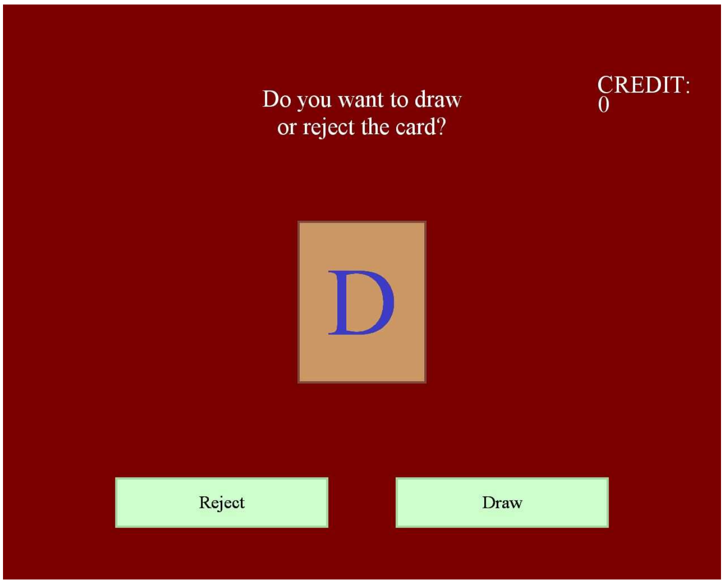
Before I introduced the use of the touchscreens in this sort of experiment, a common practice at the acoustics department was to equip experimental computers with keyboards to which some keys had been removed. Once the experiment is loaded in Presentation the software prevents the user from interacting with applications other than itself unless the escape key is pressed. One can load the experiment, remove keyboard and mouse and plug them in after it has ended, being sure that the user did not have the possibility through the touchscreen to exit from the test environment.

Graphical game interface

In the test that has been excuted the game interface presented to the participant is reported in figure 3.1(a). In the top right part of the screen a counter kept track of the total amount of virtual money earned by the subject. In the middle of the screen the presented deck (in this trial we see deck D is shown), appears and two choice buttons are displayed. Upon drawing the card from the top of the deck the result of the choice was revealed and the counter updated, as shown in figure 3.1(b). When drawing the result was shown for 1.4s, while switching had a 1s transition phase.

3.3 Test creation

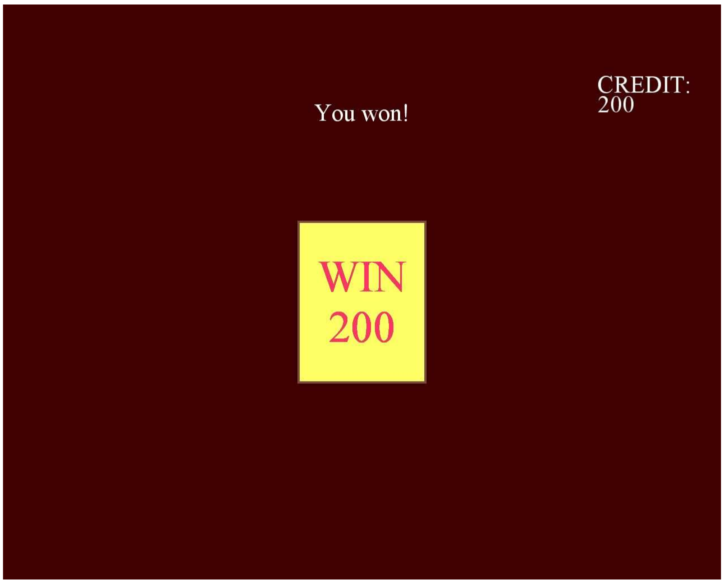
Soundscapes and IGT were put together in an experiment. As the main goal of the study was to assess whether a pleasant sound could have had a different impact on subjective decisions with respect to an unpleasant one, each participant should have perfomed tha gambling task twice, one with the pleasant sound playing in background, one with the unpleasant sound. It was interesting for us to also have an a-priori and an a-posteriori evaluation of the emotional impact of the sound: I therefore included in the test a rating phase before and after each of the two games. Before the actual game started a small tutorial provided the basic information to the subject about the task goals and the game dynamics, together with a description of the rating scales (such scales have been described in section 2.1.4). A demo of both the rating and the game mechanisms was made before the experiment started. Before each game the corresponding soundscape was played for 3 minutes without possibility of interaction with the system, then the user was asked to rate the sound and, once finished, the gamble would have started. The soundscape was looping while each game proceeded. At the end of the gamble a second rating of the soundscape was requested. For half of the participant the first sound to be played was the pleasant one, while it was the unpleasant one for the other half of them.



Do you want to draw
or reject the card?

CREDIT:
0

(a) A screenshot of the game.



You won!

CREDIT:
200

(b) A screenshot of the game outcome when drawing; rejecting the card does not give any feedback.

Figure 3.1: A sample trial of the game in which the subject draws a card from deck D

Concerning the implementation strategy Presentation, as the name suggests, is a software based on the presentation of stimuli. It has a quite simple and intuitive developing language where, though, business logic and graphic representation are mixed together and difficult to separate; the expressive power is quite limited and yet sufficient for test design purposes. The mechanism relies on the concept of trial, which basically consists in one or several stimuli, either visual or acoustic, or both. Under this philosophy I have developed a set of core functions for displaying objects on the screen that are invoked by three main modules:

- the *game module*, responsible to create instances of the IGT;
- the *rating module*, dealing with instances of the sound rating environment;
- the *slideshow module*, taking care of the introduction to game, ratings and messages between each instance of them.

The slideshow module interacts with the game and rating modules for issuing demo instances of the two. The game module invokes the rating module at the beginning and at the end of a not-demo instance to collect the participant's evaluations. Game and rating modules produce text files recording respectively game proceedings and rating results.

3.4 Test technologies and setup

3.4.1 Computers

The main requirement for the computers was the presence of the serial port for connecting the touchscreens. The computing power was a negligible aspect, as the experiment was kept simple and the number of multimedia stimuli, which are the main cause of latence for the Presentation software (see [25]), was limited. In fact only the two sound stimuli could have impacted on the performances at run-time, but the software has the possibility of loading them in memory before starting the experiment, overcoming the issue.

3.4.2 Sound cards and Headphones

The two computers were not equipped with sound cards. At the department several external audio devices were available. I used Creative EMU USB cards for reproducing the auditory stimuli. In a pretest phase I tuned the output volumes to be equivalent for both the test computers.

I used Koss headphones for the test, commonly employed at the Applied Acoustics department for teaching activities.

3.4.3 Touchscreens

The touchscreens used were 3M's M170 single touch with a maximum resolution of 1280x1024 pixels. The calibration in these devices was crucial for a correct correspondence between the buttons areas in the experiment and the screen surface. A build-in function of the vendor's software allows to perform this step with a certain ease, even if some conflicts between the Presentation software and the preexisting calibration settings imposed caution.

3.4.4 Room

The room elected for hosting the test was the lecture room at the Applied Acoustic department at Chalmers. The available technologies allowed two participants to be tested at once. A black panel was used to separate two subjects performing the task. This was done for reducing external interferences from other participants.

3.5 Test execution

The test took place at the Applied Acoustics department the days 15th, 16th, 17th, 19th, 20th, 21st, and 22nd of December of 2010. Choosing the week before Christmas granted me the availability of the lecture room at the department with the possibility of leaving the experimental setup untouched for the whole testing period.

34 people (15 males, age 28,29+-2,82) were tested during the week. The sample population was composed both by students and workers in different disciplines. Some of the subject had been contacted before for other typologies of tests at the department, while others belonged to my network of people.

All of them were aware they would have received a cinema ticket for their cooperation. They were actually rewarded with two cinema tickets for having participated. In order to enhance their focus on the task at the beginning of then experiment we said that the first ticket was granted, while the second one would have been given only for particularly good performances.

In the introduction to the experiment people were explicitly instructed to pursue the goals of maximizing the wins and minimizing the losses. In previous research (see [16]), it has been argued that such advices reduce the risk of having participants which do not take the time to evaluate the presented decks and go for random choices or regular choice patterns with deliberate avoidance of the exploration of the decks.

The average time for a participant to complete the experiment was approximately 25 minutes.

Chapter 4

Results

4.1 Data collected and clusters

4.1.1 Ratings

For each participant the following ratings were collected before and after each game and stored on a continuous line ranging from 0 to 8:

Val : perceived *valence* of the soundscape;

Act : perceived *activation*;

Sca : perceived level of *scare*;

Dem : perceived level of *demand*;

Fam : perceived level of *familiarity*.

For the statistical analysis the ratings were augmented by 1 in order to make them comparable with the IADS ratings (ranging from 1 to 9).

4.1.2 IGT

For each card presented the following data were collected:

pId : the *participant ID*;

$gNum$: the *current game* (first or second);

$pOrd$: the *sound order in the experiment* (1 if the pleasant sound was in the first game, 2 if it was in the second);

t : the *trial number* for the current game;

tc : the *selection time* (also called *reaction time*);

$deck$: the *presented deck* in the current trial;

k : the *choice made* (acceptance or switching to another deck);

win : the *card outcome* (positive for wins, negative for losses and equal to -1 when the card was not drawn).

4.1.3 Participants' basic data

For each participant these data were also collected:

asian : whether or not the subject was *asian*;

age : the *age* or the subject;

hear : whether or not the subject had *hearing impairments*;

male : whether the subject was a *male or a female.*

4.1.4 Clusters within the population

Asian vs non-Asian

Although it might sound more of a racial discrimination, in many studies (e.g. in [1]), we find significant differences between Asian (in literature mainly Japanese) and non-Asian (American), subjects' behavior. The are multiple factors which have been identified as the main actors in differentiating among cultures, such as the system of values, society structure, family hierarchies. What emerged can be summarized in the fact that Asian people are often unable to perceive a shift in their emotional state and less sensitive to external affecting events.

In the population under analysis in the second test 16 subjects out of 34 were Asian (Chinese, Singaporean, Thai), therefore taking into account as a grouping criterion this distinction appeared as a reasonable way of proceeding.

Male vs Female

This is probably the most basic division and one of the most used in many studies in literature (e.g in the studies on the IADS database, [26]).

Our population consisted of 15 males and 19 females, thus justifying a group division of this kind.

Pleasant sound order

An other division can be made between subjects who have been exposed to the pleasant sound in the first round of the game and those who had it in the second round. A difference in average responses according to whether or not the pleasant sound was played first could show us a *precedence effect.*

4.1.5 Clusters within the game: card groups

5 groups of 20 accepted cards

As every participant rejected cards in different proportion with respect to all the others we had to find a way of grouping the choices to have comparable indexes. The idea was to divide the total amount of cards presented in each game into 5 sets with a total of 20 accepted cards each. Therefore every 20 accepted cards we computed the partial statistics to be used in the analyses by taking into account all the cards presented up to the 20th accepted one.

2 groups of 20 accepted cards

After we grouped the cards of a trial as described in the previous section, for some analyses we focused on only a couple of adjacent groups, the second and the third. This was done under suggestion of the Swedish supervisors to better understand the behavior in the central phase of the game.

2 groups of 25 accepted cards

Similarly with what has ben done with groups of 20, we also clustered cards 25-by-25 and we kept only the 2 initial groups to investigate differences, in any, between the very beginning of the game and the central part. In a second moment we repeated the grouping discarding the first 2 accepted cards of every deck, which corresponded to eliminate the cards that were presented in fixed order at the beginning of the game as described in section 3.2.1.

4.2 Absolute ratings of the soundscapes

Table 4.1 shows mean and standard deviation of the ratings of the soundscapes. A one-sample t-test was run on the ratings of valence and activation before the game for both the soundscapes comparing them to the mean values obtained by averaging the ratings of the IADS sounds in them. The results are summarized in table 4.2, where we can see that no significant differences where found. This result can be seen as a *supporting evidence of **Hypothesis 3** validity*.

4.2.1 Demand insight

Tests were run using General linear models to assess dependence of the perceived level of demand and the soundscape being played. A full factorial design with the 2 sounds and the 2 ratings before and after the game was made. The results showed a significant interaction of sound and rating occurrence ($F = 5.754$, $Sig. = 0.022$, $\eta_P^2 = 0.148$). Figure 4.1 shows the

Statistics of Ratings

		Pleasant soundscape		**Unpleasant soundscape**	
Dimension		Mean	Std. Dev.	Mean	Std. Dev.
Valence	before	7,42	1,33	3,3	1,44
	after	6,36	1,73	3,24	1,59
Activation	before	5,63	1,67	6,55	1,32
	after	5,5	1,91	5,7	1,8
Familiarity	before	7,16	1,9	5,92	2,62
	after	7,13	1,79	6,5	2,18
Scare	before	1,96	1,68	5,45	2,26
	after	2,55	2,09	4,43	2,25
Demand	before	4,96	2,01	6,24	1,79
	after	5,07	1,96	5,22	2,15

Table 4.1: Basic statistics for the ratings in the test

Pleasant soundscape			
Valence	IADS mean	t	Sig.
	7,4	0,09	0,93
Activation	IADS mean	t	Sig.
	5,27	1,27	0,21
Unpleasant soundscape			
Valence	IADS mean	t	Sig.
	3,02	1,13	0,27
Activation	IADS mean	t	Sig.
	6,97	-1,87	0,07

Table 4.2: t-tests on valence and activation of the soundscapes

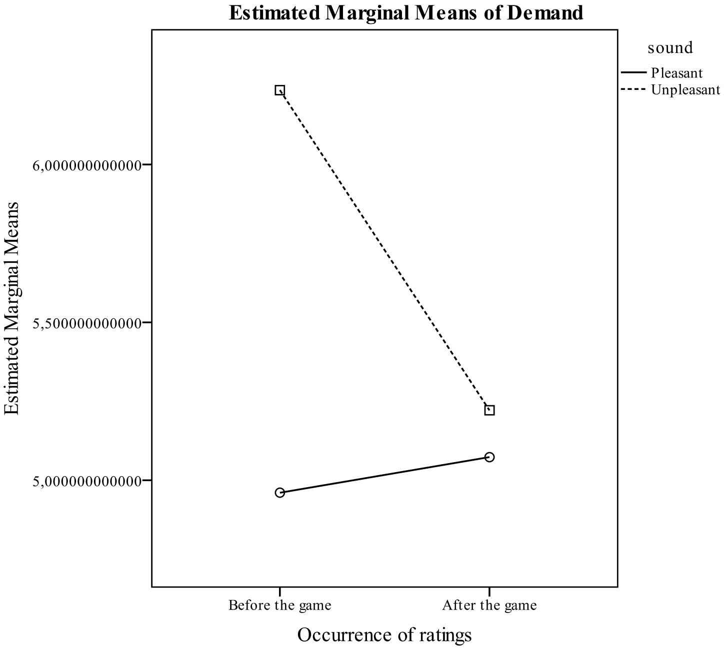


Figure 4.1: Demand variation from before to after the IGT.

demand value for the positive and negative soundscapes. As can be seen, while the rating for the pleasant sound does not change significantly before and after the IGT, the unpleasant sound decreases and almost reaches the same value as the pleasant setting after the game.

4.2.2 Scare and Familiarity

The two dimensions Familiarity and Scare are not deeply analyzed, as they are regarded more as accessorial dimensions to help subject's self assessment of emotional content of the soundscapes.

4.3 IGT

4.3.1 Dependent variables for the tests

The second is the *percentage of rejections from bad decks.* It is computed as the ratio between the number of rejected cards from decks A and B and the total number of times these decks were presented in the group.

$$\text{rej\_if\_bad} = \frac{\text{number of rejected cards from decks A and B}}{\text{total number of cards presented from decks A and B}} \tag{4.3.1}$$

Other variables were computer from the data which could have given more flexibility in the analysis by detaching from the nature of decks.

The first one is the *percentage of drawn cards in the group.* This measure is given by the sum of all the accepted cards in a group divided by the total number of cards presented in the group.

$$\text{chosen\_norm} = \frac{\text{number of accepted cards in the group}}{\text{total number of cards presented}} \tag{4.3.2}$$

4.3.2 Hypotheses restated

Recalling Hypothesis 1:

> ***Hypothesis 1:***
> ***Positive mood-inducing sounds increase risk aversion in decision-making.***

Detailing what reported by Blanchette,

> Losses are perceived even more negatively by happy participants than they are by control participants. Risk aversion may thus stem from an increased motivation to avoid losses that are perceived as being more consequential by participants in positive states... [p.16 [3]]

We can now associate this hypothesis to a suitable variable to be tested. The chosen quantity is the rejection ratio of high losses decks (rej_if_highloss): if a positive mood increases risk aversion in thus enhances loss aversion this should reflect in a tendency to avoid cards which potentially could lead to greater losses than others. The first hypothesis then becomes:

> ***Hypothesis 1:***
> ***Positive mood-inducing sounds increase the rejection ratio of decks with higher losses.***

Hypothesis 2 states:

> ***Hypothesis 2:***
> ***Negative mood-inducing sounds increase risk taking in decision-making.***

An increase in risk taking can be related to a higher acceptance ratio of cards, regardless of the deck. We than rephrase hypothesis 2:

> ***Hypothesis 2:***
> ***Negative mood-inducing sounds increase the acceptance ratio of cards.***

4.3.3 Tests on the hypotheses of the IGT

For hypothesis 1 we have rejection ratios of high losses decks for every group of 20 accepted cards, both for positive and negative soundscapes. We can then use a general linear model with the 2 different sounds and the 5 groups of cards as factors ($2 \cdot 5$ full factorial design). The software used for the computations is SPSS, which can provide us with several output indexes such as the observed power η_P^2 and generate plots of the interactions between factors. The results of the simulations did not show any significant interaction of sound on the value of the rejection ratios ($F = 0.949$, $Sig. = 0.337$, $\eta_P^2 = 0.028$). Also the interaction due to sound and card group together was not significant ($F = 0.596$, $Sig. = 0.666$, $\eta_P^2 = 0.018$). The only significant factor was the deck group ($F = 4.553$, $Sig. = 0.002$, $\eta_P^2 = 0.121$). In figure 4.2 we can see the evolution of the rejection ratio across the card groups. Simulations with only card groups 2 and 3 and with 2 card groups of 25 cards (teh first 2 card groups in the trial), led to similiar, not significant results.

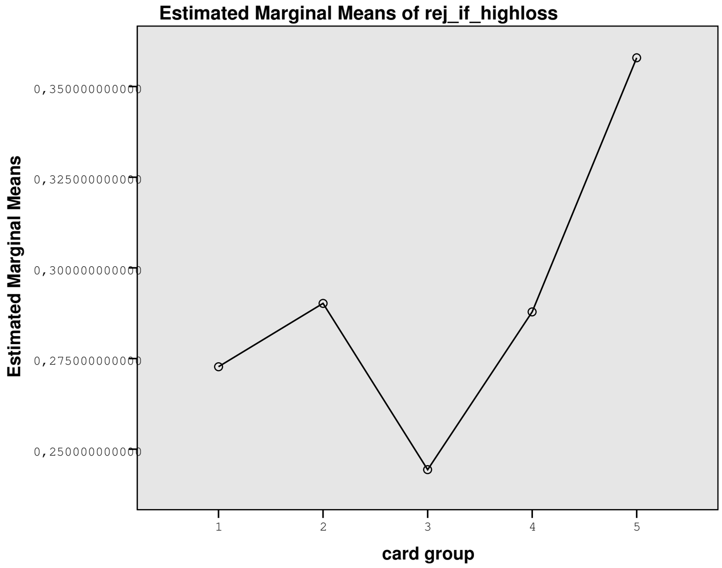


Figure 4.2: Evolution of the rejection ratio of high loss decks through card groups.

For hypothesis 2 we have ratios of chosen cards over the total number of cards displayed in each of the 5 card groups. A new general linear model analysis is performed with a setup analogous to the one for testing hypothesis 1. Also for hypothesis 2 the results show no significant influence of sound on choice ratio ($F = 0.734$, $Sig. = 0.398$, $\eta_P^2 = 0.022$). Also the combined interaction of sound and card group is not significant ($F = 0.225$, $Sig. =$

0.924, $\eta_P^2 = 0.007$). Again we have a significant dependence of the choice ratio on the card group ($F = 4.960$, $Sig. = 0.001$, $\eta_P^2 = 0.131$).

Chapter 5

Discussion

5.1 Absolute ratings of the soundscapes

As shown in table 4.2, no significant difference was found between the average ratings of the IADS sounds used in each soundscape creation and the subjects ratigns of the soundscapes themselves. This validates Hypothesis 3 that states:

> ***Hypothesis 3:***
> ***The emotional content of a soundscape can be predicted by averaging the emotional content of sounds previously categorized according to specific affective ratings mechanisms.***

We could infer from this result that for conveying a certain emotional content with a synthesized soundscape we can approximate the impact it will have by averaging the emotional content of its constituting key sounds.

For what concerns the level of demand we found that it decreases for the negative sound, wihle it remains stable for the positive one. This could mean that a negative mood-inducing sound could make the subject more focused on the task, progressively disregarding the sound itself. This finding would go against the changing state hypothesis, which states a crossmodal disruptive effect taht was not supported by evidence here.

5.2 IGT Hypothesis test

No supporting evidence was found for hypotheses 1 and 2. The reasons can be several:

- design flaws which prevented us from avoiding bias in the response behavior of the participants;
- wrong choice of the dependent variables for the hypothesis test;

- number of subjects insufficient to have a good database of responses to be analyzed, forcing us to avoid as much as possible the removal of outliers.

5.3 Conclusion

The goal of the present study was to open the way to the investigation of the effects of sound stimuli on learning processes and specifically in how sound can affect our decision-making strategies. What we assessed was a consistence between the soundscape affective ratings and the affective ratings of their key constituent sound samples. It was not possible to find evidences of direct effects of sound on decision-making, although further investigation is advisable.

5.3.1 Future development

The acceptance of hypothesis 3 has given a good reference point for a future, deeper investigation of the possibilities and limits of this design principle.

For what concerns the game, some research has been done in order to formalize the problem in mathematical terms. An introduction to the method is provided in appendix B.

Chapter 6

Conclusions

Appendix A

Dependent variables for the IGT

In this short appendix all the dependent variables utilized for tests on the IGT data 4 are presented with a short explanation.

"Good" versus "Bad"

Three different indicators were devised for the analysis.
The first one is the *percentage of accepted cards from good decks.* This measure is given by the sum of all the accepted cards coming from decks C and D divided by the number of times decks C and D where displayed

$$\text{acc\_if\_good} = \frac{\text{number of accepted cards from decks C and D}}{\text{total number of cards presented from decks C and D}} \tag{A.0.1}$$

The second is the *percentage of rejections from bad decks.* It is computed as the ratio between the number of rejected cards from decks A and B and the total number of times these decks were presented in the group.

$$\text{rej\_if\_bad} = \frac{\text{number of rejected cards from decks A and B}}{\text{total number of cards presented from decks A and B}} \tag{A.0.2}$$

The last equation represents the *normalized proportion of accepted cards from good decks.* It is given by the ratio of the acc_if_good variable over the sum of acc_if_good and the percentage of accepted cards from bad decks (A and B).

$$\text{acc\_good\_norm} = \frac{\text{acc\_if\_good}}{\text{acc\_if\_good} + (1 - \text{rej\_if\_bad})} \tag{A.0.3}$$

"High wins" versus "Low wins"

Two different indicators were devised for the analysis.
The first one is the *percentage of accepted cards from high winning decks.*

This measure is given by the sum of all the accepted cards coming from decks B and D divided by the number of times decks C and D where displayed

$$\text{acc\_if\_highwin} = \frac{\text{number of accepted cards from decks B and D}}{\text{total number of cards presented from decks B and D}} \tag{A.0.4}$$

The second is the *percentage of rejections from low wins decks.* It is computed as the ratio between the number of rejected cards from decks A and C and the total number of times these decks were presented in the group.

$$\text{rej\_if\_lowwin} = \frac{\text{number of rejected cards from decks A and C}}{\text{total number of cards presented from decks A and C}} \tag{A.0.5}$$

“High losses” versus “Low losses”

Two different indicators were devised for the analysis.
The first one is the *percentage of accepted cards from low losses decks.* This measure is given by the sum of all the accepted cards coming from decks A and D divided by the number of times decks C and D where displayed

$$\text{acc\_if\_lowloss} = \frac{\text{number of accepted cards from decks A and D}}{\text{total number of cards presented from decks A and D}} \tag{A.0.6}$$

The second is the *percentage of rejections from high losing decks.* It is computed as the ratio between the number of rejected cards from decks B and C and the total number of times these decks were presented in the group.

$$\text{rej\_if\_highloss} = \frac{\text{number of rejected cards from decks B and C}}{\text{total number of cards presented from decks B and C}} \tag{A.0.7}$$

A.0.2 Deck independent variables

Other variables were computer from the data which could have given more flexibility in the analysis by detaching from the nature of decks.
The first one is the *percentage of drawn cards in the group.* This measure is given by the sum of all the accepted cards in a group divided by the total number of cards presented in the group.

$$\text{chosen\_norm} = \frac{\text{number of accepted cards in the group}}{\text{total number of cards presented}} \tag{A.0.8}$$

The second equation represents the *average response time in a group.* It takes into account response times regardless of the choice made in the trials.

$$\text{choice\_time\_avg} = \frac{\text{sum of response times}}{\text{total number of cards presented}} \tag{A.0.9}$$

The last two equations try to take into account the past experiences influence. They are respectively the *percentage of cards that have been rejected when the previous one was a winning one* and the *percentage of cards that have been rejected when the previous one was a losing one*:

$$\text{rej\_after\_pos} = \frac{\text{number of rejected cards when the previous one was a win}}{\text{total number of winning cards}} \tag{A.0.10}$$

$$\text{rej\_after\_neg} = \frac{\text{number of rejected cards when the previous one was a loss}}{\text{total number of losing cards}} \tag{A.0.11}$$

Appendix B

Modeling the learning problem

In this appendix a quick review of some possible models for the present study are proposed.

B.1 Introduction

In section 2.1 I have been addressing the main principles on which the decision-making task will be grounded. How to translate them into a plausible mathematical framework to develop an analysis?
In literature some attempts can be found to formalize learning mechanisms. In interesting studies by Yechiam and Busemeyer (see [27] and [28]), several models for experience-based decision-making were compared. They focused on two different tasks, the *original IGT* and the *Go/No-Go Discrimination task*.

The former was described in section 2.1.2. The Go/No-Go tasks is another psychological tool that was employed in assessing, among others, differences between normal subjects and criminals. It consists of a random sequence of 10 numbers, 5 of which are given a positive outcome and the remaining 5 a negative one. The subject has to decide whether to accept or reject the number which is presented. Upon acceptance the corresponding reward (or punishment) was given. Rejecting the number was not associated to any event.

The structure of the Go/No-Go task is formally similiar to the version of the IGT I chose for the experiment:

- each number shown in a Go/No-Go trial can be associated with a different deck label of the modified IGT trial;

- the acceptance/rejection of the Go/No-Go reflects in the acceptance/deck-switching option of the IGT.

B.2 The proposed models

In forming the models I merged some concepts from the formal analysis of the original IGT and the Go/No-Go task, adapting them to the peculiarities of the modified version of the IGT I have used in the experiment. In general we can say that learning models are based on understanding how an individual makes decisions based on past outcomes. In our case what we need first is therefore a measure of how much wins and losses of a specific trial impact on the subject. This is achieved in the model thanks to a *utility function* $u(t)$, where the variable t is the trial we are evaluating. In a second time we have to create estimates of how the perceived quality of a deck evolves through trials. This helps us in forecasting choices according to the experience gained. An *expectancy function* $E_j(t)$ is devised for each deck j and each trial t. We will see how this function can be related to past values of the expectancy itself and to the present utility. Once the expectancies are updated we can use them as parameters of a *choice rule* computing the probability of drawing or switching deck.

B.2.1 The utility function

In order to account for the positive and negative outcomes a simple and yet robust utility function $u(t)$ which has been validated in several studies (CIATCIATCAITCAI), assumes a linear dependence from the amounts won and lost in each trial:

$$u(t) = (1 - w) \cdot Win(t) - w \cdot Loss(t) \tag{B.2.1}$$

The term w is a number between 0 and 1 that we can term *attention*; it is a weight of negative outcomes in the trial t with respect to positive ones. $Win(t)$ is the positive quantity won at trial t, while $Loss(t)$ is the lost quantity in case of a negative result from the choice.

B.2.2 Update of expectations

In the generic equation of the current expectation for the deck j $E_j(t)$ (which is updated after the choice has been made), we still assume a linear dependence from two quantities: the previous value of expectation and the utility. In formula:

$$E_j(t) = \alpha_{jt} E_j(t-1) + \beta_{jt} u(t) \tag{B.2.2}$$

The terms α_{jt} and β_{jt}, the coefficients of the equation that in general depend on both the current deck and the trial, are defined according to different assumptions. Here we take into account three possible models of how the

Interference EV model (IEV)

In this model we do not differentiate between decks, as it is assumed that people's reaction only occurs due to the outcome, regardless of the deck it is associated with.

$$E(t) = E(t-1) + a \cdot k(t) \cdot [u(t) - E(t-1)] \tag{B.2.3}$$

The term a denotes the weight of the prediction error $[u(t) - E_j(t-1)]$ and k is 1 when I draw the card from the deck, 0 if I switch. When I switch deck the expectation is not updated, as $k(t)$ is 0.

Decay EV model

Expectation of each deck is updated every trial.

$$E(t) = a \cdot E(t-1) + k(t) \cdot u(t) \tag{B.2.4}$$

If I switch the presented deck at trial t (that is $k(t) = 0$), then we only have the term $a \cdot E(t-1)$, where a is a decay factor: the more we proceed in the game, the more we try to insert new information in working memory. Therefore it is reasonable to assume that oldest cards displayed will be progressively forgotten, regardless of the deck they belong.

Interference CD model (ICD)

Expectations deck j are updated only when I am presented with that deck and I choose an action.

$$E_j(t) = E_j(t-1) + a \cdot \delta_j(t) \cdot k(t) \cdot [u(t) - E_j(t-1)] \tag{B.2.5}$$

The term a denotes the weight of the prediction error $[u(t) - E_j(t-1)]$ and $\delta_j(t)$ is a dummy variable which is 1 when the deck displayed int trial t is deck j, 0 otherwise. The term k is 1 when I draw the card from the deck, 0 if I switch; this tells us that when deck j appear and I do not draw, E_j does not get updated.

Decay CD model (DCD)

Expectations of each deck are updated every trial.

$$E_j(t) = a \cdot E_j(t-1) + \delta_j(t) \cdot k(t) \cdot u(t) \tag{B.2.6}$$

If the presented deck is not the j-th then we only have the term $a \cdot E_j(t-1)$, where a is a decay factor a $k(t)$ the choice made, as in equation B.2.4. $\delta_j(t)$ again tells me which deck is now presented.

Interference MCV model (IMCV)

This model tries to integrate the IEV and ICD perspectives giving a weight to cues coming from not just the deck related to the expectation, but also from other ones. In any case only when the card is chosen I have an update of the expectations.

$$E_j(t) = E_j(t-1) + a \cdot k(t) \cdot \{\gamma \cdot [1-\delta_j(t)] + (1-\gamma) \cdot \delta_j(t)\} \cdot [u(t) - E_j(t-1)] \tag{B.2.7}$$

Decay MCV model (DMCV)

This model tries to integrate the IEV and ICD perspectives giving a weight to cues coming from not just the deck related to the expectation, but also from other ones. In this way at every trial we have an update of the expectation.

$$E_j(t) = a \cdot E_j(t-1) + k(t) \cdot \{\gamma \cdot [1-\delta_j(t)] + (1-\gamma) \cdot \delta_j(t)\} \cdot u(t) \tag{B.2.8}$$

B.2.3 Choice rule

In the EV model we have one expectancy.

$$P[G(t+1) = k_0] = \frac{e^{\theta(t) \cdot E_{k_0}(t)}}{\sum_{k=0}^{1} e^{\theta(t) \cdot E_k(t)}} \tag{B.2.9}$$

In the CD and MCV models we have different expectancies for each deck, therefore:

$$P[G_j(t+1) = k_0] = \frac{e^{\theta(t) \cdot E_{j,k_0}(t)}}{\sum_{k=0}^{1} e^{\theta(t) \cdot E_{j,k}(t)}} \tag{B.2.10}$$

We can make different assumptions about the parameter c. Two main streams are outlined in the following subsections.

Trial-dependent consistency

The term $\theta(t)$ is not constant, but changes on time, accounting for an *exploratory phase* at the beginning of the game. In formula.

$$\theta(t) = (t/10)^c \tag{B.2.11}$$

Trial-independent consistency

The term $\theta(t)$ is constant through trials.

$$\theta(t) = 3^c - 1 \tag{B.2.12}$$

Bibliography

[1] Lisa Feldman Barrett, Batja Mesquita, Kevin N Ochsner, and James J Gross. The experience of emotion. *Annual Review of Psychology*, 58(1):373–403, 2007.

[2] Fredrik Hagmar. Emotional response to sound Influence of spatial determinants. Master's thesis, Chalmers University of Technology, 2010.

[3] Isabelle Blanchette and Anne Richards. The influence of affect on higher level cognition: A review of research on interpretation, judgement, decision making and reasoning. *Cognition & Emotion*, 24(4):561–595, 2010.

[4] A Damasio. *Descartes Error*. Putnam, 1994.

[5] A Bechara, A R Damasio, H Damasio, and S W Anderson. Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, 50(1-3):7–15, 1994.

[6] Robert Schleicher, Shiva Sundaram, and Julia Seebode. Assessing audio clips on affective and semantic level to improve general applicability. In *DAGA*, 2010.

[7] Lotte F Van Dillen and Sander L Koole. Clearing the mind: a working memory model of distraction from negative mood. *Emotion Washington Dc*, 7(4):715–723, 2007.

[8] T Meiser and Karl Christoph Klauer. Working memory and changing-state hypothesis. *Journal of Experimental PsychologyLearning Memory Cognition*, 25(5):1272–1299, 1999.

[9] Alan D Baddeley. Working memory: looking back and looking forward. *October*, 4(10):829–839, 2003.

[10] A Baddeley. The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, 4(11):417–423, 2000.

[11] Ana Tajadura-Jimnez, Aleksander Vljame, Erkin Asutay, and Daniel Vstfjll. Embodied auditory perception: the emotional impact of approaching and receding sound sources. *Emotion Washington Dc*, 10(2):216–229, 2010.

[12] Penny Bergman, Anders Sköld, Daniel Västfjäll, and Niklas Fransson. Perceptual and emotional categorization of sound. *The Journal of the Acoustical Society of America*, 126(6):3156–67, December 2009.

[13] Roberto Lucchetti. *Di duelli, schacchi e dilemmi: la teoria matematica dei giochi*. Bruno Mondadori, 2002.

[14] Danielle Barry and Nancy M Petry. Predictors of decision-making on the Iowa Gambling Task: independent effects of lifetime history of substance use disorders and performance on the Trail Making Test. *Brain and Cognition*, 66(3):243–252, 2008.

[15] DDG Schneider-Bakos and MAP Parente. A cultural study on decision making: performance differences on the Iowa gambling task between selected groups of Brazilians and Americans, 2010.

[16] E Peters and P Slovic. The springs of action: Affective and analytical information processing in choice. *Personality and Social Psychology Bulletin*, 26(12):1465–1475, 2000.

[17] M M Bradley and Peter J Lang. Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1):49–59, 1994.

[18] R L Hodes, E W Cook, and Peter J Lang. Individual differences in autonomic response: conditioned association or conditioned fear? *Psychophysiology*, 22(5):545–560, 1985.

[19] Jens Gisselgrd, Julia Uddn, Martin Ingvar, and Karl Magnus Petersson. Disruption of order information by irrelevant items: a serial recognition paradigm. *Acta Psychologica*, 124(3):356–369, 2007.

[20] F Firestone. The phase difference and amplitude ratio at the ears due to a source of pure tone. *Journal of the Acoustical Society of America*, 2:260–270, 1930.

[21] Ircam. LISTEN HRTF DATABASE, 2003.

[22] Effect size - Wikiversity. “`http://en.wikiversity.org/wiki/Effect_size`”, 2011.

[23] Jean Stockard, Robert M O’Brien, and Ellen Peters. The use of mixed models in a modified Iowa Gambling Task and a prisoners dilemma game. *Judgment and Decision Making*, 2(1):9–22, 2007.

[24] Shiva Sundaram, Robert Schleicher, and Julia Seebode. CLUSTERING AUDIO CLIPS BY CONTEXT-FREE DESCRIPTION AND AFFECTIVE RATINGS. In *EUSIPCO*, pages 472–476, 2010.

[25] Neurobehavioral Systems. Presentation - Documentation.

[26] Ryan a. Stevenson and Thomas W. James. Affective auditory stimuli: Characterization of the International Affective Digitized Sounds (IADS) by discrete emotional categories. *Behavior Research Methods*, 40(1):315–321, February 2008.

[27] Eldad Yechiam and Jerome R Busemeyer. Comparison of basic assumptions embedded in learning models for experience-based decision making. *Psychonomic bulletin review*, 12(3):387–402, 2005.

[28] Eldad Yechiam, Jackson Goodnight, John E Bates, Jerome R Busemeyer, Kenneth A Dodge, Gregory S Pettit, and Joseph P Newman. A formal cognitive model of the go / no-go discrimination task : Evaluation and implications. *Psychological Assessment*, 18(3):239 –249, 2006.

[29] Sø ren Bech and Nick Zacharov. *Perceptual Audio EvaluationTheory, Method and Application*, volume 0. John Wiley and Sons, 2006.

[30] Lee A. Becker. Measures of Effect Size (Strength of Association), 1999.

[31] C Castelfranchi and M Miceli. The Cognitive-Motivational Compound of Emotional Experience. *Emotion Review*, 1(3):223–231, 2009.

[32] Eswar Damaraju, Yang-Ming Huang, Lisa Feldman Barrett, and Luiz Pessoa. Affective learning enhances activity and functional connectivity in early visual cortex. *Neuropsychologia*, 47(12):2480–7, 2009.

[33] Gesine Dreisbach. How positive affect modulates cognitive control: the costs and benefits of reduced maintenance capability. *Brain and Cognition*, 60(1):11–19, 2006.

[34] Renata M Heilman, Liviu G Crisan, Daniel Houser, Mircea Miclea, and Andrei C Miu. Emotion regulation and decision making under risk and uncertainty. *Emotion Washington Dc*, 10(2):257–265, 2010.

[35] Alice M Isen. An Influence of Positive Affect on Decision Making in Complex Situations: Theoretical Issues With Practical Implications. *Journal of Consumer Psychology*, 11(2):75–85, 2001.

[36] Mendel Kleiner. *Audio Technology and Acoustics.* Goteborg, 2008.

[37] B C J Moore. *An Introduction to the Psychology of Hearing*, volume 3. Academic Press, 2003.

[38] John A. Rice. *Mathematical Statistics and Data Analysis.* Duxbury Press, 2nd edition, 1995.

[39] Lotte F Van Dillen, Dirk J Heslenfeld, and Sander L Koole. Tuning down the emotional brain: an fMRI study of the effects of cognitive load on the processing of affective images. *NeuroImage*, 45(4):1212–1219, 2009.

[40] Lotte Van Dillen and Sander Koole. How automatic is automatic vigilance? The role of working memory in attentional interference of negative information. *Cognition & Emotion*, 23(6):1106–1117, 2008.

[41] ANOVA: power and size. "`http://www.linguistics.ucla.edu/faciliti/facilities/statistics/power.htm`", 2011.

[42] J A Easterbrook. The effect of emotion on cue utilization and the organization of behavior. *Psychological Review*, 66(3):183–201, 1959.

[43] C E Y Evans, K Kemish, and O H Turnbull. Paradoxical effects of education on the Iowa Gambling Task. *Brain and Cognition*, 54(3):240–244, 2009.

[44] Gary W. Evans, Karen Mata Allen, Richard Tafalla, and Tiffany O'Meara. Multiple stressors: Performance, psychophysiological and affective responses. *Journal of Environmental Psychology*, 16:147–154, 1996.

[45] Patrick Gomez and Brigitta Danuser. Affective and physiological responses to environmental noises and music. *International Journal of Psychophysiology*, 53(2):91–103, 2004.

[46] Seong-Hoon Kang Kang and Sung-Han Kim Kim. Realistic Audio Teleconferencing using Binaural and Auralization Techniques. *ETRI Journal*, 18(1):41–51, April 1996.

[47] WJ Macken, N Mosdell, and D M Jones. Explaining the irrelevant-sound effect: Temporal distinctiveness or changing state? *Journal Of Experimental Psychology. Learning Memory And Cognition*, 25(3):810–814, 1999.

[48] Kiheiji Nishida and Yuichiro Kanazawa. On the variance-stabilizing multivariate nonparametric regression estimation. 2010.

[49] R F Simons, B H Detenber, T M Roedema, and J E Reiss. Emotion processing in three systems: the medium and the message. *Psychophysiology*, 36(5):619–627, 1999.

[50] Joshua A Weller, Irwin P Levin, and Antoine Bechara. Do individual differences in Iowa Gambling Task performance predict adaptive decision making for risky gains and losses? *Journal of Clinical and Experimental Neuropsychology*, 32(2):141–150, 2010.

[51] Soundscape of European Cities and Landscapes. "`http://www.soundscape-cost.org/`", 2011.

[52] Antoine Bechara, H Damasio, and A R Damasio. Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, 10(3):295–307, 2000.

[53] A R Damasio. Descartes' error revisited. *Journal of the History of Neuroscience*, 10(2):192–4., 2001.

[54] Daren Zhang, Xiaochu Zhang, Xiwen Sun, Zhihao Li, Zhaoxin Wang, Sheng He, and Xiaoping Hu. Cross-modal temporal order memory for auditory digits and visual locations: an fmri study. *Human Brain Mapping*, 22(4):280–289, 2004.

[55] Susan E Gathercole and Tracy Packiam Alloway. An introduction to working memory. In *Working Memory and Learning: A Practical Guide for Teachers*, chapter 1, page 144. SAGE Publications Ltd, 2008.

[56] M M Bradley and Peter J Lang. Affective reactions to acoustic stimuli. *Psychophysiology*, 37(2):204–215, 2000.

[57] Ruud Wetzels, Joachim Vandekerckhove, Francis Tuerlinckx, and Eric-Jan Wagenmakers. Bayesian parameter estimation in the expectancy valence model of the iowa gambling task. *Journal of Mathematical Psychology*, 54(1):14–27, 2010.

[58] I Myung. Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*, 47(1):90–100, 2003.

[59] Jerome R Busemeyer and In Jae Myung. An adaptive approach to human decision making: Learning theory, decision theory, and human performance. *Journal of Experimental Psychology: General*, 121(2):177–194, 1992.

[60] StatSoft. General Linear Models (GLM), 2011.